

# Investment Analysis (Análisis de Inversiones)

## Apuntes del Material Docente

Szabolcs István Blazsek-Ayala

### Table of contents

Financial data	1
Financial risk	9
Portfolio risk	14
Portfolio theory	21
Forecasting security prices	34
Factor models	42
Dynamic models of volatility	52
Jumps in security prices: Markov switching models	63
Extreme values in security returns: value at risk	66
Portfolio selection based on cointegration	74

## Financial data

## CLASS STRUCTURE

1. **Asset returns**
2. **Portfolio returns**
3. **Distributional properties of returns**
  - density function and histogram
  - moments (mean, variance, skewness, kurtosis)
  - multivariate returns (covariance matrix)
4. **Financial data sources**

### Asset returns

- Financial econometrics data mostly involves returns, instead of prices, of assets.
- The two main reasons for this choice are:
- (1) For average investors, return of an asset is a **complete** and **scale-free** summary of the investment opportunity.
- (2) Return series are easier to handle than price series because the former have more attractive statistical properties.

### Asset returns

- Let  $P_t$  be the price of an asset at time index  $t$ .
- We introduce two alternative ways to compute asset returns:
  1. **One-period simple return**,  $R_t$
  2. **Continuously compounded return or log return**,  $r_t$

### Asset returns

#### 1. One-period simple return

- Holding the asset for one period from date  $t-1$  to date  $t$  would result in the simple return:

$$R_t = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}}$$

### Asset returns

#### Continuous compounding

- Before introducing continuously compounded return, we discuss the idea of continuous compounding.

## Asset returns

### Continuous compounding

- Assume that the interest rate of a bank deposit is 10% per annum and the initial deposit is \$1.
- If the bank pays interest once a year, then the net value of the deposit becomes  $\$1(1+0.1) = \$1.1$  one year later.

## Asset returns

### Continuous compounding

- If the bank pays interest semi-annually, the 6-month interest rate is  $10\%/2 = 5\%$  and the net value is  $\$1(1+0.1/2)^2 = \$1.1025$  after the first year.
- In general, if the bank pays interest  $m$  times a year, then the interest rate for each payment is  $10\%/m$  and the net value of the deposit becomes  $\$1(1+0.1/m)^m$  one year later.

## Asset returns

### Continuous compounding

- The following table gives the results for some commonly used time intervals on a deposit of \$1.00 with interest rate 10% per annum.

## Asset returns

	m	interest rate per period	net value
Annual	1	10.00%	1.10000
Semiannual	2	5.00%	1.10250
Quarterly	4	2.50%	1.10381
Monthly	12	0.83%	1.10471
Weekly	52	0.19%	1.10506
Daily	365	0.03%	1.10516
Continuously	$\infty$		1.10517

## Asset returns

### Continuous compounding

- The effect of compounding is clearly seen in the table.
- In particular, the net value approaches \$1.10517, which is obtained by  $\exp(0.1)$  and referred to as the result of continuous compounding.

## Asset returns

### 2. Continuously compounded return

- The continuously compounded return or log return is defined as

$$r_t = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}}$$

## Asset returns

### 2. Continuously compounded return

- The main reason why the use of log return is preferred to the simple return in financial econometrics is because the statistical properties of log returns are more tractable.
- Therefore, in this course, we focus on the log returns.

## Portfolio returns

- Consider a portfolio of  $N$  assets.
- Let  $p$  be a portfolio that places weight  $w_i$  of total portfolio value on asset  $i$  when the portfolio is created.
- Then the simple return and log return of  $p$  at time  $t$  are given by

$$R_{pt} = \sum_{i=1}^N w_i R_{it} \quad r_{pt} \simeq \sum_{i=1}^N w_i r_{it}$$

## Portfolio returns

- The error of the approximation of the log return formula is large when the change of prices of individual assets is large.
- Typically, on the financial markets the daily price changes are such that the error of the log return approximation formula is relatively little.

## Dividend payment

- If an asset pays dividends periodically, we must modify the definitions of asset returns.
- Let  $D_t$  be the dividend payment of an asset between dates  $t-1$  and  $t$  and  $P_t$  be the price of the asset at the end of period  $t$ . Thus, dividend is not included in  $P_t$ .

## Dividend payment

- Then, the simple net return and continuously compounded return at time  $t$  become

$$R_t = \frac{P_t + D_t}{P_{t-1}} - 1$$

$$r_t = \ln(P_t + D_t) - \ln(P_{t-1})$$

## Distributional properties of returns

- A main characteristic of financial econometrics is that asset returns a-priori are uncertain.
- This feature of asset returns is captured by modeling asset returns mathematically as **random variables**.
- In this section, we introduce some concepts of random variables to be used later in the course.

### Distributional properties of returns

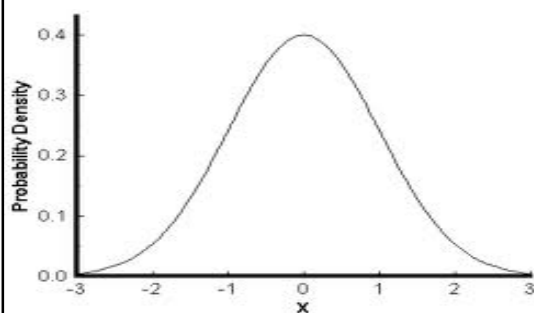
- To study asset returns, we begin with their distributional properties.
- The objective is to understand the behavior of the returns across assets and over time.
- Consider a collection of  $i=1, \dots, N$  assets held for  $T$  time periods, say  $t=1, \dots, T$ .
- For each asset  $i$ , we study its log return at time  $t$ ,  $r_{it}$ .

### Distributional properties of returns

#### Density function of returns

- The properties of a distribution can be characterized graphically by the **density function**.
- The density function informs about the probabilities of alternative values of the random variable.
- See the following graph:

### Distributional properties of returns



### Distributional properties of returns

#### Density function of returns

- Interpretation: The higher the value of the density function  $f(X)$ , the higher the probability of observing the value  $X$ .
- We can see on the graph that  $f(X)$  corresponds to a symmetric distribution, centered at the value zero.
- Nevertheless, in the reality the distribution of returns is not necessarily symmetric and centered at zero.

### Distributional properties of returns

#### Density function of returns

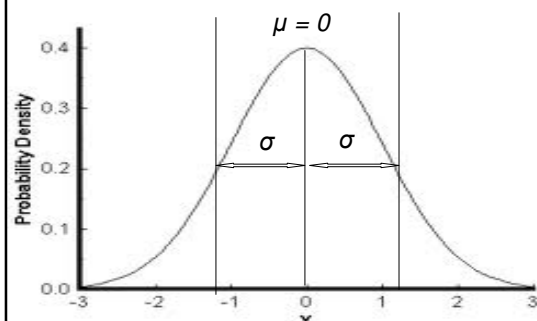
- A reference probability distribution used frequently in financial econometrics is the so-called **normal distribution** denoted as

$$X \sim N(\mu, \sigma^2)$$

where  $\mu$  and  $\sigma$  denote the **mean** (or **average**) and the **standard deviation** of  $X$ .

(See the next figure.)

### Distributional properties of returns



## Distributional properties of returns

### Density function of returns

- The density  $f(X)$  of a normal distribution is given by

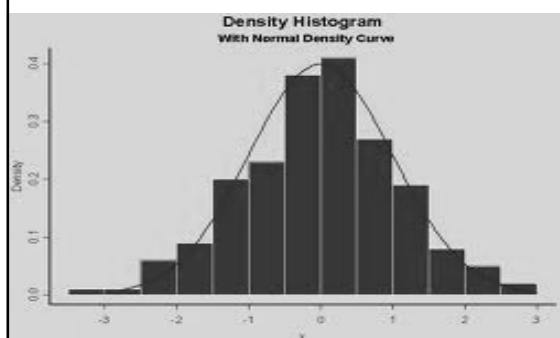
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

## Distributional properties of returns

### Histogram of returns

- The density function for real data can be represented by the histogram.
- The histogram is graphed over a set of **classes** over the data set.
- The histogram presents the **number of observations** counted in each of the classes.
- **Exercise.** See the following figure for an example of a histogram.

## Histogram and normal density



## Distributional properties of returns

### Moments of a random variable

- The random variable,  $X$  can be characterized by its **moments**.
- We introduce four moments and discuss why they are useful to understand the behavior of the random variable.

## Distributional properties of returns

### Moments of a random variable

(1) **Mean** or **expectation** of  $X$  denoted by  $\mu_X$

- The mean of a random variable is interpreted as the **average** value of  $X$ .
- Typically,  $\mu_X$  of financial asset returns is very close to the zero value.
- When market participants **forecast** future returns, they are interested in estimating  $\mu_X$ .

## Distributional properties of returns

### Moments of a random variable

(2) **Variance** of  $X$  denoted by  $(\sigma_X)^2$

- We note that the **standard deviation** of  $X$  is the square root of the variance and is denoted by  $\sigma_X$ .
- When  $X$  is the asset return, the standard deviation is also called **volatility**.
- The volatility is interpreted as the dispersion of the random variable and in practice it is often used as a measure of financial risk of investments.

## Distributional properties of returns

### Moments of a random variable

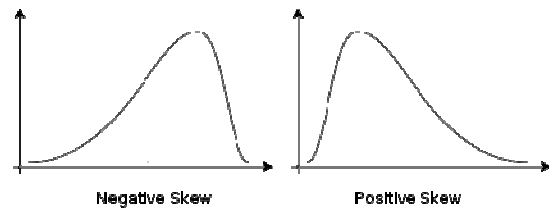
#### (3) Skewness denoted by $S(X)$

- The skewness informs about the **asymmetry** of the distribution of  $X$ .
- If  $S(X) < 0$  then the mass of the distribution is concentrated on the right.
- If  $S(X) > 0$  then the mass of the distribution is concentrated on the left.
- If  $S(X) = 0$  then the distribution is symmetric.

## Distributional properties of returns

### Moments of a random variable

#### (3) Skewness denoted by $S(X)$



## Distributional properties of returns

### Moments of a random variable

#### (4) Kurtosis denoted by $K(X)$

- Kurtosis summarizes **tail thickness** of the distribution.
- When a distribution has high tail thickness this means that **extreme values** of  $X$  occur with relatively high probability.
- Investors on a financial market are very interested in the occurrence of **extreme returns**, for example market crashes.

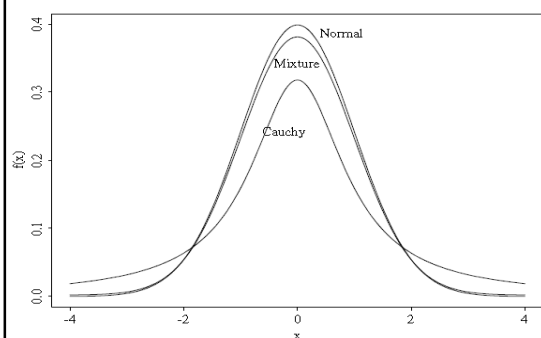
## Distributional properties of returns

### Moments of a random variable

#### (4) Kurtosis denoted by $K(X)$

- The following figure shows three distributions with different kurtosis:

## Distributional properties of returns



## Distributional properties of returns

### Moments of a random variable

#### (4) Kurtosis denoted by $K(X)$

- The  $K(X) = 0$  of a normal random variable, which is used as a reference value.
- A distribution with  $K(X) > 0$  is said to have heavy tails, implying that the distribution puts more mass on the tails of its support than a normal distribution does.
- In practice, this means that a random sample from such a distribution tends to contain more extreme values.

## Distributional properties of returns

### Estimation of moments

- In application, the moments can be estimated from the data.
- Let  $\{x_1, \dots, x_T\}$  be a random sample of  $X$  with  $T$  observations.
- We can compute the moments in Excel using its statistical functions in a very simple way.

## Distributional properties of returns

- In practice, there exist more complicated statistical models for the moments (e.g., mean, variance) of asset returns.
- This exercise is an introduction for moment estimation.

## Distributional properties of returns

### Multivariate returns

- In application, we are often interested in the statistical properties of the relationships between different assets' returns besides the moments mentioned previously (mean, variance, skewness, kurtosis) because the single assets form portfolios whose properties depend on these relationships.
- Let  $r_t = (r_{1t}, \dots, r_{Nt})'$  be a vector of log returns of  $N$  assets at time  $t$ .

## Distributional properties of returns

### Multivariate returns

- The pair wise relationships among different returns are described by the so-called variance-covariance matrix,  $\Sigma$ .
- When we consider a portfolio of  $N$  assets, the  $\Sigma$  is an  $N \times N$  quadratic matrix.
- The diagonal of  $\Sigma$  contains the variances, while the off diagonal elements are the covariances denoted  $\text{cov}(r_{it}, r_{jt})$ .

## Distributional properties of returns

### Multivariate returns

- The covariance  $\text{cov}(r_{it}, r_{jt})$  can be interpreted computing the so-called correlation coefficient between  $r_{it}$  and  $r_{jt}$ :

$$\text{Correlation coefficient} = \rho_{ij} = \frac{\text{Cov}(r_{it}, r_{jt})}{\sigma_i \sigma_j}$$

## Distributional properties of returns

### Multivariate returns

$$-1 \leq \rho_{ij} \leq 1$$

- If  $\rho_{ij} = 0$  then the two returns are not correlated.
- If  $\rho_{ij} < 0$  then the two returns are negatively correlated, i.e. they are moving in the opposite direction.
- If  $\rho_{ij} > 0$  then the two returns are positively correlated, i.e. they are moving in the same direction.



## Distributional properties of returns

### Multivariate returns

- There are three properties of  $\Sigma$ :
  - (1)  $\Sigma$  is quadratic
  - (2)  $\Sigma$  is symmetric
  - (3)  $\Sigma$  is positive semi-definite

## Distributional properties of returns

### Multivariate returns

- (1) The first property says that the number of columns and rows is equal.
- (2) The second property says that  $\text{cov}(r_{it}, r_{jt}) = \text{cov}(r_{jt}, r_{it})$ . Thus, we can interchange the rows and the columns, i.e. transpose, without changing  $\Sigma$ .

## Distributional properties of returns

### Multivariate returns

- (3) The third property is analogous to the fact that the variance is a positive number.
- The definiteness of  $\Sigma$  can be verified by computing the determinant  $D$  of all principal sub-matrices of  $\Sigma$ .
- The  $\Sigma$  is a positive semi-definite matrix if the determinant of all principal sub-matrices is non-negative.
- The determinant can be computed in Excel by the MDETERM function.

## Data sources

- There is a large amount of financial market data bases available for practitioners.
- Maybe the two most famous are Bloomberg and Reuters.
- In this course, we shall download data from the following web-sites:
  - <http://www.google.com/finance>
  - <http://finance.yahoo.com/>
- From these sites, one can download to Excel daily / weekly / monthly price data.

## Financial risk

## INVESTMENT PROCESS

Investment process =

- (1) Security and market analysis +
- (2) Formation of an optimal portfolio of assets

## INVESTMENT PROCESS

- The objective of “(1) *Security and market analysis*” is to assess the **expected-return**  $\mu$  and **risk**  $\sigma$  attributes of the entire set of investment assets.
  - In the following slides, we relate  $\mu$  and  $\sigma$ .
  - See the topics of **forecasting**, **factor models** for  $\mu$  and the topics of **dynamic volatility models** and **value at risk** for  $\sigma$  for additional information.

## INVESTMENT PROCESS

- The purpose of “(2) *Formation of an optimal portfolio of assets*” is the determination of the best risk-return opportunities available from feasible investment portfolios.
  - See the topic of **portfolio theory** for this problem.

## INVESTMENT PROCESS

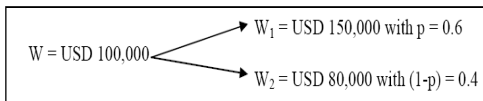
- (1) The **expected-return**  $\mu$  and **risk**  $\sigma$  attributes of the entire set of investment assets can be analyzed at least in two ways:
  - (a) Compare the **risk premium** of the risky assets.
  - (b) Evaluate the  $\mu$  and  $\sigma$  attributes of the risky asset using the **utility function**.
- In the followings, we introduce these two concepts.

(a) Risk premium

### Example

- Consider the following one-period investment where the initial wealth, USD 100,000 is invested in a risky asset.
- In this example, the presence of risk means that more than one outcome is possible.
- This can be represented by the following tree of outcomes:

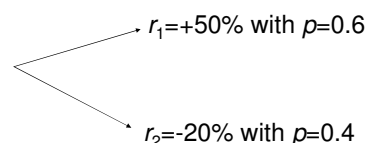
(a) Risk premium



We can see that with probability  $p = 0.6$  the favourable outcome will occur, leading to final wealth USD 150,000. However, with probability  $(1-p) = 0.4$  the less favourable outcome will occur with a final wealth of USD 80,000.

(a) Risk premium

- In terms of **simple returns**, the tree can be written as follows:



(a) Risk premium

- In order to evaluate this investment, first we look at the descriptive statistics of the distribution of the simple return:

**(1) Expected return,  $\mu$**

$$\begin{aligned}\mu &= r_1 \times p + r_2 \times (1-p) \\ &= (+50\%) \times 0.6 + (-20\%) \times 0.4 = 22\%\end{aligned}$$

- Therefore, the expected return of the investment is 22%.

(a) Risk premium

**(2) Variance of return,  $\sigma^2$**

$$\begin{aligned}\sigma^2 &= p[r_1 - \mu]^2 + (1-p)[r_2 - \mu]^2 = \\ &= 0.6[50\% - 22\%]^2 + 0.4[-20\% - 22\%]^2 = \\ &= 0.0996\end{aligned}$$

- **Standard deviation of the return,  $\sigma$**  =  $\sigma = 31.6\%$

- In finance, it is common to define financial risk as the standard deviation of the return.

(a) Risk premium

- Question: Is the  $\sigma = 31.6\%$  risk high or low?
- One way to answer this question is to compare the return of the risk investment with the return of alternative investments.
- A typically chosen benchmark for the alternative investment is the government Treasury-bill (T-bill), which is considered as a **risk-free security**.

(a) Risk premium

- In our example, consider the T-bill as alternative investment.
- Suppose that the T-bill offers a rate of return of  $r_f = 5\%$ .

#### (a) Risk premium

- We can define the **risk premium** of the risky investment as the difference between the expected return on the risky investment and the return on the risk-free investment:

- **RISK PREMIUM** =  
 $= E(r_i) - r_f = 22\% - 5\% = 17\%$

#### (b) Utility function

- An alternative way to assess the  $\mu$  and  $\sigma$  attributes of the risk asset is to apply the utility function, which quantifies the utility obtained by the investor holding the risky asset.
- Higher utility values are assigned to portfolios with more attractive risk-return profiles.
- The utility function has two arguments: expected return  $\mu$  and risk  $\sigma$ .

#### (b) Utility function

- A commonly used utility function is:

$$U = \mu - A \sigma^2$$

where the  $A$  real number is the **risk aversion coefficient**.

- The value of  $A$  determines different types of investors:
  - Risk averse investor
  - Risk neutral investor
  - Risk lover investor

#### RISK AVERSE INVESTORS

- A **risk averse** investor penalizes the expected rate of return of a risky portfolio to account for the risk involved.
- An example of the risk-return utility function is the following:

$$U = \mu - A \sigma^2$$

where  $A > 0$ .

#### RISK NEUTRAL INVESTORS

- **Risk neutral** investors judge risky investments based on only the expected rate of return of the investment.
- The level of risk is irrelevant to this investor that is there is no penalization for risk.
- An example of the utility function for a risk-neutral investors is the following:

$$U = \mu$$

- In other words, the risk aversion index  $A = 0$  for a risk neutral investor.

#### RISK LOVER INVESTORS

- **Risk lover** investors are willing to invest in risky projects:
- They adjust upward the utility function for higher risk. Thus, take into account the fun of the investment's risk.
- For a risk lover investor an example of the utility function is the following:

$$U = \mu - A \sigma^2$$

where  $A < 0$  is an index of the investor's risk loving (or negative risk aversion).

### MEAN-VARIANCE CRITERION

- In the reality, and empirically proven fact is that most investors tend to be **risk averse**.
- That is they penalize expected return by risk.
- Risk averse investors ( $A > 0$ ) can choose between two alternative risky investments using the so-called **mean-variance criterion**.

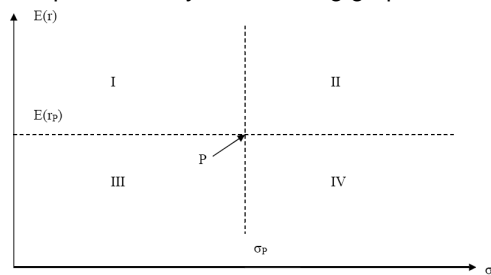
### MEAN-VARIANCE CRITERION

- Formally, we can state the **mean-variance criterion** as follows:
- Investment  $A$  dominates investment  $B$  if

$$E(r_A) \geq E(r_B) \text{ and } \sigma_A \leq \sigma_B$$

### MEAN-VARIANCE CRITERION

The mean-variance criterion can be represented by the following graph:



### MEAN-VARIANCE CRITERION

- In the middle of the figure portfolio  $P$  is presented with expected return  $E(r_P)$  and standard deviation  $\sigma_P$ .
- A risk averse investor prefers  $P$  to any portfolio in quadrant IV because  $P$  has higher expected return and lower risk than any investment in IV.
- Moreover, any portfolio in quadrant I is preferable to  $P$  as they have higher expected return and lower risk.

### MEAN-VARIANCE CRITERION

- What can be said about quadrants II and III?
- In order to compare the portfolios of these quadrants, we need more information about the exact nature of the investor's risk aversion.

### MEAN-VARIANCE CRITERION

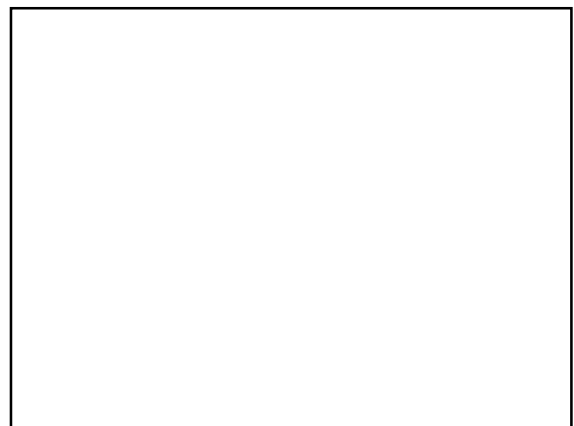
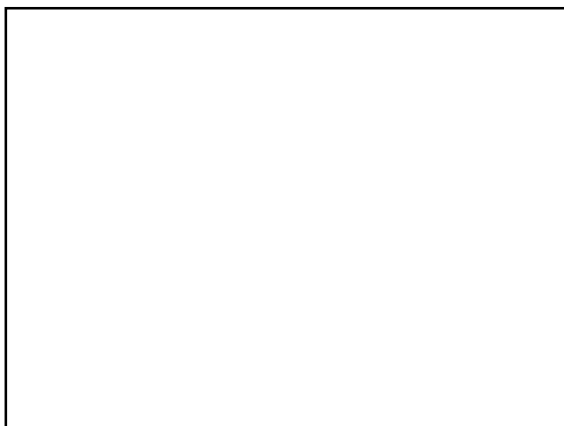
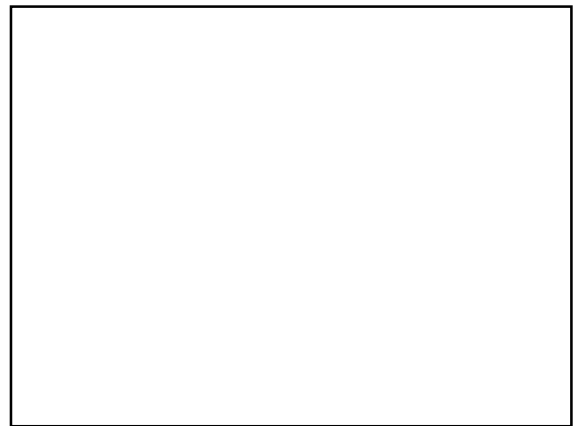
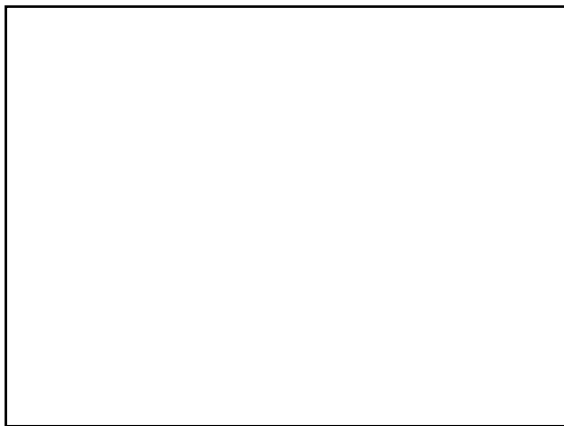
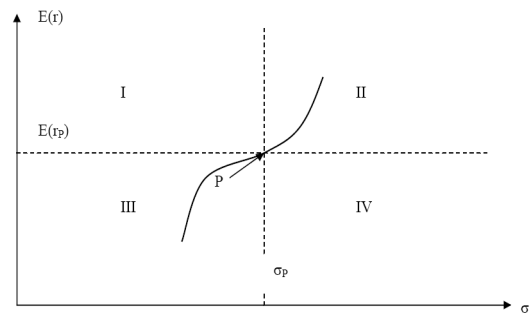
- Suppose an investor identifies all portfolios that are equally attractive as portfolio  $P$ . Starting at  $P$ , an increase in standard deviation must be compensated by an increase in expected return.
- Investors will be equally attracted to portfolios with high risk and high expected returns compared with other portfolios with lower risk but lower expected returns.

## MEAN-VARIANCE CRITERION

### Indifference curve:

- These equally preferred portfolios will lie on a curve in the mean-standard deviation graph that connects all portfolio points with the same utility value.
- This curve is called **indifference curve**.

## MEAN-VARIANCE CRITERION Indifference curve



## Portfolio risk

### ASSET RISK VS PORTFOLIO RISK

- In this section, we focus on the expected return and risk attributes of a **portfolio**.
- A portfolio is a set of  $i=1, \dots, N$  single assets.
- In the reality, investors allocate their funds in many assets that form a portfolio.
- We will show that the overall risk of a portfolio may be smaller than the risk of the single assets included in the portfolio.

### ASSET RISK VS PORTFOLIO RISK

- This may be due to two different reasons:
  - (1) **Hedging**: investing in an asset with a payoff pattern that offsets the portfolio's exposure to a particular source of risk.
  - (2) **Diversification**: investments are made in a wide variety of assets so that the exposure of risk of any particular security is limited.

### PORTFOLIO RETURN

Expected return of a portfolio

- **Proposition** (Expected return of a portfolio): The expected return of a portfolio is the weighted average of individual asset expected returns where the weights,  $w_i$  for  $i = 1, \dots, N$ , are the asset proportions in the portfolio of  $N$  assets:

$$E[r_P] = \sum_{i=1}^N w_i E[r_i]$$

### PORTFOLIO RISK

Portfolio risk for a risk-free and a risky asset

- **Proposition** (Standard deviation of a portfolio of a risky and a risk-free asset):
- The portfolio standard deviation,  $\sigma_P$  equals the risky asset's standard deviation,  $\sigma$  multiplied by the portfolio proportion,  $w$  invested in the risky asset:

$$\sigma_P = |w\sigma|.$$

### PORTFOLIO RISK

Portfolio risk for two risky assets

- **Proposition** (Variance of a portfolio of two risky assets):
- The portfolio variance  $\sigma_P^2$  of two risky assets with variances  $\sigma_1^2$  and  $\sigma_2^2$ , respectively and weights  $w_1$  and  $w_2=1-w_1$ , respectively is given by:

$$\sigma_P^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \text{cov}(r_1, r_2)$$

## PORTFOLIO RISK

### Portfolio risk for two risky assets

- Remark:

○ Covariance =  $\text{cov}(r_1, r_2) = \sigma_1 \sigma_2 \rho_{12}$  where  $-1 \leq \rho_{12} \leq 1$  denotes the **correlation coefficient**.

- Thus, the previous formula can also be written as:

$$\sigma_p^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_1 \sigma_2 \rho_{12}$$

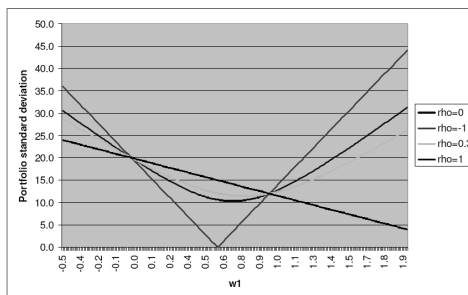
## EXAMPLE

### Portfolio risk for two risky assets

- On the following slide, we present the portfolio variance of a 2-asset portfolio as a function of the weight of the first asset for alternative values of  $\rho_{12}$ :

## EXAMPLE

### Portfolio risk for two risky assets



## EXAMPLE

### Portfolio risk for two risky assets

- Consider four scenarios for the correlation coefficient.
- Assume that the standard deviation of each asset is the same in all scenarios.
- Compute the portfolio variance for different values of  $w_1$  and  $w_2$ .

## INTERPRETATION OF THE FIGURE

- In the previous figure we can see that it is possible that  $w_1 < 0$  and  $w_2 > 1$ .

○ In these cases, we **go short** of asset 1 and invest the money obtained in asset 2.

- It is also possible that  $w_1 > 1$  and  $w_2 < 0$ .

○ In these situations, we **go short** of asset 2 and invest the cash obtained in asset 1.

## SHORT SELLING

- What does it mean to “go short”?
- We “**go short**” or “**short sell**” a financial product if we sell the product without having it.
- This means that at the time of selling the product, after borrowing from other investor, we obtain cash.
- This also means that in the future we have to buy back the same product and give back to the other investor.



### SHORT SELLING

- Why is it interesting for an investor to short sell a product?
- Because “going short” of a product is just the opposite investment strategy as “buying” a product:
- When we buy a product we speculate on increasing price. LONG POSITION
- When short sell a product we speculate on decreasing price. SHORT POSITION

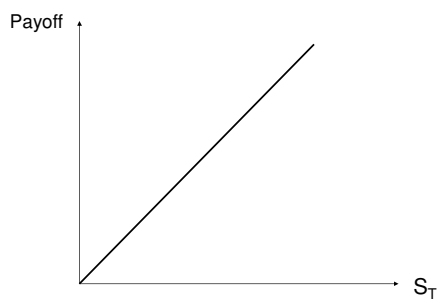
### LONG AND SHORT POSITIONS

- In the following figures, the payoff and profit of the long and short positions in a stock are presented.

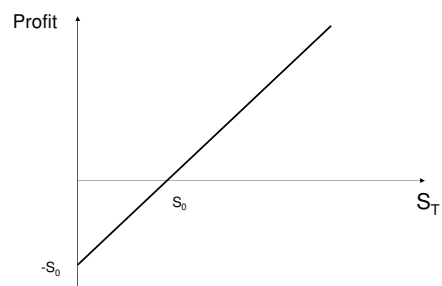
Notation:

- $S_0$  = the price of the financial product at time  $t=0$  when the buy or ‘short sell’ happens.
- $S_T$  = the price of the financial product at time  $t=T$  when the corresponding sell or ‘buy back’ happens.

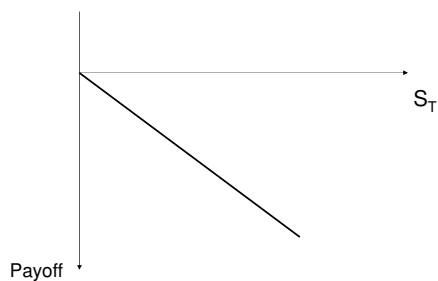
#### LONG POSITION PAYOFF at time $t=T$



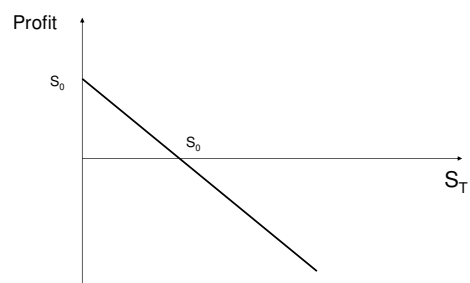
#### LONG POSITION PROFIT at time $t=T$



#### SHORT POSITION PAYOFF at time $t=T$



#### SHORT POSITION PROFIT at time $t=T$



## PORTFOLIO RISK

Portfolio with large number of assets

- We extend the portfolio variance for an arbitrary number of assets.
- **Proposition** (Variance of a portfolio of several risky assets):
- Denote the  $N \times 1$  vector of portfolio weights by  $w$  and let  $\Sigma$  be an  $N \times N$  variance-covariance matrix of  $N$  asset returns. Then, the variance of the portfolio of  $N$  assets is given by:

$$\sigma_p^2 = w' \Sigma w$$

## PROPERTIES OF $\Sigma$

- There are three properties of  $\Sigma$ :
  - (1)  $\Sigma$  is quadratic
  - (2)  $\Sigma$  is symmetric
  - (3)  $\Sigma$  is positive semi-definite.
- **Positive semi-definite** means that
$$\sigma_p^2 = w' \Sigma w \geq 0$$
for any real values of  $w$  (even for  $w < 0$ )
- The consequence of (3) is that the portfolio variance is non negative.

## HOW TO CHECK THAT $\Sigma$ IS POSITIVE SEMI-DEFINITE?

- The definiteness of  $\Sigma$  can be verified by computing the determinant  $D$  of the principal sub-matrices.
- There is a function for this in Excel: MDETERM(sub-matrix).
- **Proposition.** A matrix is positive semi-definite if and only if the determinant of all principal sub-matrices is non-negative.

## EXAMPLE

Portfolio variance for three risky assets

- Consider the case of three assets:
- Doing simple algebra, from  $w' \Sigma w$  we can express the portfolio variance of three assets as follows:
$$\begin{aligned} \sigma_p^2 = & w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + w_3^2 \sigma_3^2 \\ & + 2w_1 w_2 \text{cov}(r_1, r_2) + 2w_1 w_3 \text{cov}(r_1, r_3) \\ & + 2w_2 w_3 \text{cov}(r_2, r_3) \end{aligned}$$

## EXAMPLE

Portfolio variance for three risky assets

- Consider four scenarios for the correlation coefficient.
- Assume that the standard deviation of each asset is the same in all scenarios.
- Suppose a particular value for  $w_3$
- Compute the portfolio variance for different values of  $w_1$  and  $w_2$ .

Portfolio variance for portfolios of thousands of assets

- In the previous slides of portfolio variance, we computed the variance small portfolios that included 2 or 3 assets only.
- However, in the reality, banks have portfolios with thousands of different assets.
- In these situations, how to compute the variance of a portfolio?

### SOLUTION

- Previously, we used the explicit solution for the formula  
Portfolio variance =  $w' \Sigma w$
- For portfolios where there are few assets this solution has a relatively simple form and we can type it into Excel to compute  $\sigma_p^2$ .
- See for example the 2 and 3 asset cases:

### SOLUTION FOR FEW ASSETS

- 2 asset case ( $N=2$ ):  
$$\sigma_p^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \text{COV}(r_1, r_2)$$

○ There are 4 terms.
- 3 asset case ( $N=3$ ):  
$$\sigma_p^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + w_3^2 \sigma_3^2 + 2w_1 w_2 \text{COV}(r_1, r_2) + 2w_1 w_3 \text{COV}(r_1, r_3) + 2w_2 w_3 \text{COV}(r_2, r_3)$$

○ There are 9 terms.

### PROBLEM FOR MANY ASSETS

- We can see in the previous equations that the number of terms is  $N^2$  where  $N$  is the number of assets in the portfolio.
- For example, imagine that a bank has 5,000 assets in its portfolio and wants to compute the risk of that portfolio. Then,  $(5,000)^2 = \mathbf{25 \text{ million}}$  terms in the equation.
- In these cases, it is not possible to derive the formula explicitly due to the large number of terms in the equation.

### SOLUTION FOR MANY ASSETS

- If  $N$  is large we need to compute the  
Portfolio variance =  $w' \Sigma w$   
formula by matrix multiplication in the computer.
- We can do this in Excel using the **MMULT(matrix\_left, matrix\_right)** function.
- In the following slides, we present how to use this function.

### EXAMPLE FOR 5 ASSETS

- For example, when  $N=5$  then there are 25 terms in the formula so it is complicated to derive and type in the computer all terms.
- Therefore, we use the MMULT function to compute portfolio risk.

### STEP 1

- We need to choose the  $5 \times 1$  vector of portfolio weights:  
$$w = (w_1, w_2, w_3, w_4, w_5)'$$
- Choose weights such that  
$$(w_1 + w_2 + w_3 + w_4 + w_5) = 1$$

## STEP 2

- We need to choose (or in practice: estimate) the parameters of the variance-covariance matrix  $\Sigma$ .
- $N = 5$  implies that there are 5 standard deviations and  $(5 \times 5 - 5) / 2$  correlation coefficients to choose.

## STEP 3

- Once we have the initial parameters of the covariance matrix we have to compute the elements of  $\Sigma$  as follows:
- The diagonal of  $\Sigma$  contains the variances:  $\sigma_i^2$  with  $i = 1, \dots, 5$ .
- The off-diagonal elements of  $\Sigma$  are the covariances:  $\sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$  with  $i = 1, \dots, 5$  and  $j = 1, \dots, 5$ .

## STEP 4

- Then, we have to check if the determinant of all principal sub-matrices is non-negative to ensure that  $\Sigma$  is a positive semi-definite matrix:
- Use MDETERM(sub-matrix).

## STEP 5

- Now we have everything to compute the Portfolio variance =  $w' \Sigma w$  formula.
- We do this in two steps because in the MMULT function we can only multiply two matrices:

## STEP 5

- Step 5a: compute  $Z = w' \Sigma$   
(As  $w'$  is  $1 \times 5$  and  $\Sigma$  is  $5 \times 5$ , therefore,  $Z$  will be a  $1 \times 5$  vector.)
- Step 5b: compute  $w' \Sigma w = Z w$   
(As  $Z$  is  $1 \times 5$  and  $w$  is  $5 \times 1$ , therefore,  $Z w$  will be a  $1 \times 1$  matrix.)

## STEP 5

Some details of the use of MMULT:

- (1) Before calling the MMULT function we always have to select the cells of the solution.
- For example, if we multiply a  $1 \times 5$  with a  $5 \times 5$  matrix then the solution will be a  $1 \times 5$  matrix. So we have to select a  $1 \times 5$  matrix of empty cells first.

### STEP 5

- (2) Then, we initiate the MMULT function.
- In the two arguments, we select the left matrix and the right matrix.
- (3) After this, we initiate the computation of the MMULT function pushing the **CTRL+SHIFT+ENTER** buttons in the computer.

### STEP 6

- Finally, we can take the square-root of the portfolio variance to get the portfolio standard deviation.

### EXAMPLE

Portfolio variance for five risky assets

- Consider only one scenario for the variance-covariance matrix and the portfolio weights and compute the portfolio standard deviation.

## Portfolio theory

### THE MARKOWITZ PORTFOLIO SELECTION MODEL

- The portfolio selection models to be presented was developed by Harry Markowitz in the 1950s.
- In the followings:
  - First, we present the idea of **diversification** of financial portfolios and show that it may reduce overall risk.
  - Second, we solve the problem of **portfolio optimization** in two steps to get the optimal portfolio weights.

## PORTFOLIO DIVERSIFICATION

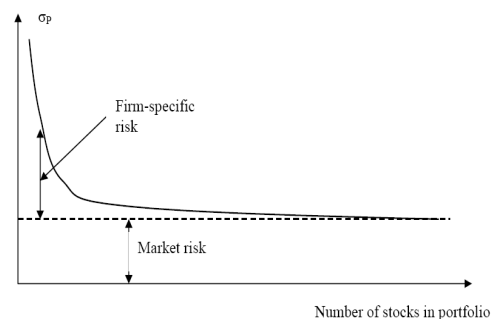
### DIVERSIFICATION

- We begin with the discussion at a general level.
- We present how **diversification** can reduce the variance (or risk) of portfolio returns.
- Diversification means the inclusion of additional (different) risky assets into the original risky portfolio.

### DIVERSIFICATION

- The following figure demonstrates the evolution of portfolio risk as a function of the number of **different** stocks included in the portfolio:

### DIVERSIFICATION



### DIVERSIFICATION

- The diversification reduces all **firm-specific risks** due to the so-called **insurance principle**.
- The reason is that with all risk sources independent, and with the portfolio spread across many securities, the exposure to any particular source of risk is reduced to a negligible level.

### DIVERSIFICATION

- However, when common sources of risk affect all firms, even extensive diversification cannot eliminate risk.
- The risk that remains even after extensive diversification is called **market risk**.

### DIVERSIFICATION

- There are different names for firm-specific risk and for market risk:
  - (1) firm-specific risk =
    - unique risk =
    - non-systematic risk =
    - diversifiable risk
  - (2) market risk =
    - systematic risk =
    - non-diversifiable risk

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- In the followings, we construct risky portfolios that provide the **lowest** possible risk for any given level of expected return.
- We prove how diversification may reduce the variance of the portfolio of **two risky assets** compared to the two individual risky assets on their own.

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- Let the sub-index 1 denote the first and 2 denote the second risky asset.
- The portfolio variance is given by
$$\sigma_p^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_1 \sigma_2 \rho$$
where  $\rho$  denotes the correlation coefficient of returns between the two risky assets.

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- As  $w_1$  and  $w_2=(1-w_1)$  influence the portfolio variance, the investor would be interested in the question of **which value of weight  $w_1$  minimizes the risk of the portfolio** for given  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  values?

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- We solve the following minimization problem:

$$\min_{w_1} \sigma_p^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 \sigma_1 \sigma_2 \rho$$

$$\min_{w_1} \sigma_p^2 = w_1^2 \sigma_1^2 + (1-w_1)^2 \sigma_2^2 + 2w_1(1-w_1) \sigma_1 \sigma_2 \rho$$

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- Take derivative with respect to  $w_1$  and equal zero the derivative:

$$2w_1 \sigma_1^2 - 2 \sigma_2^2 + 2w_1 \sigma_2^2 + 2 \sigma_1 \sigma_2 \rho - 4w_1 \sigma_1 \sigma_2 \rho = 0$$

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- Solve this equation for  $w_1$ :

$$w_1^* = [\sigma_2^2 - \text{cov}(r_1, r_2)] / [\sigma_1^2 + \sigma_2^2 - \text{cov}(r_1, r_2)]$$

where  $\text{cov}(r_1, r_2) = \sigma_1 \sigma_2 \rho$  and

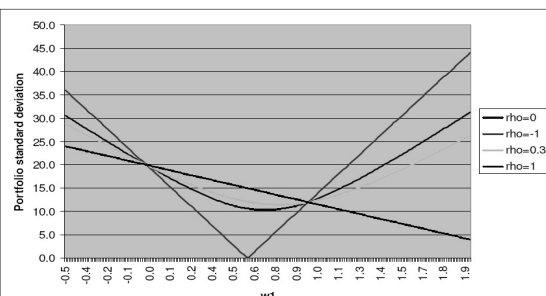
$$w_2^* = 1 - w_1^*$$

- The portfolio with weights  $w_1^*$  and  $w_2^*$  define the **minimum variance portfolio**.

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- We can also demonstrate the minimum variance portfolio on the following figure of portfolio standard deviation as a function of  $w_1$ :

### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio



### PORTFOLIOS OF TWO RISKY ASSETS Minimum variance portfolio

- In the figure, we present  $w_1 < 0$ ,  $w_2 > 1$  investment where we go short of asset 1 and invest the obtained money in asset 2.
- In addition, we also present  $w_1 > 1$ ,  $w_2 < 0$  position where we go short of asset 2 and invest the obtained cash in asset 1.



## PORTFOLIOS OF TWO RISKY ASSETS

### Minimum variance portfolio

- We derived the values of  $w_1$  and  $w_2$ , which implies the minimum variance portfolio.
- Nevertheless, besides the level of portfolio risk, the portfolio manager is also interested in the expected return of the portfolio managed.
- In the followings, we review this, more complex, problem in two steps.

## PORTFOLIO OPTIMIZATION

## PORTFOLIO THEORY

- Portfolio managers seek to achieve the best possible trade-off between risk and return.
- Suppose that the manager has to choose an optimal combination of **several risky assets and one risk-free asset**.
- We structure the portfolio manager's decision problem into the following two steps:

### STEP 1: Construct the **optimal risky portfolio** from the risky assets.

- Step (1a) **Asset allocation decision**: the choice about the distribution of the risky asset classes (stocks, bonds, real estate, foreign assets, etc).
- Step (1b) **Security selection decision**: the choice of which particular securities to hold within each asset class.
- In other words, the optimal choice from the available **risky securities** is made in step 1 to set up the **optimal risky portfolio**.

### STEP 2: Construct the **optimal complete portfolio** from the optimal risky portfolio and the risk-free asset.

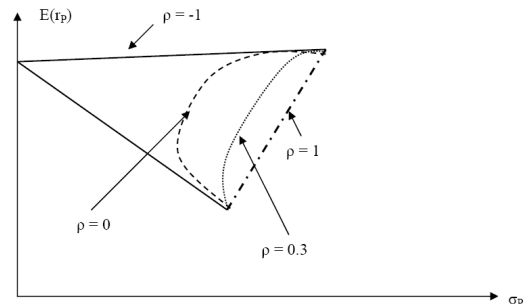
- Step (2) **Capital allocation decision**: the choice of the proportion of the **risk-free asset** and the **optimal risky portfolio**.

### STEP 1: OPTIMAL RISKY PORTFOLIO

### Portfolio opportunity set

- In the following figure, we present the expected return of the portfolio  $E(r_p)$  as a function of the standard deviation of the portfolio return  $\sigma_p$  for a portfolio of two risky assets for different portfolio weights  $w_1$  and  $w_2$ .
- This figure presents the **portfolio opportunity set**.

### Portfolio opportunity set



### Portfolio opportunity set

- The portfolio opportunity set shows the combination of expected return and standard deviation of all portfolios that can be constructed from the two available risky assets by choosing different portfolio weights  $w_1$  and  $w_2$ .
- The different lines on the figure correspond to the different value of the correlation coefficient,  $\rho$ .

### Portfolio opportunity set

- The straight line corresponding to the  $\rho = 1$  case shows that there is no benefit from diversification when perfect correlation of the two risky assets is observed.

### Portfolio opportunity set

- The lowest value of the correlation coefficient is  $\rho = -1$ . When this case happens then the investor has the opportunity of creating a perfectly hedged position by choosing the portfolio weights as follows:  

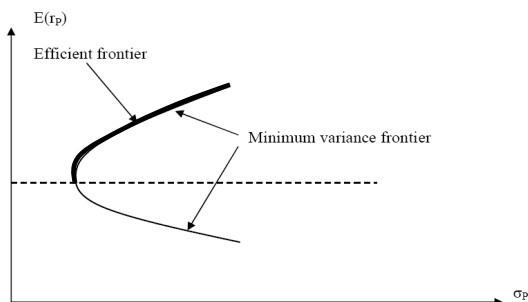
$$w_1 = \sigma_2 / (\sigma_1 + \sigma_2) \quad \text{and}$$

$$w_2 = \sigma_1 / (\sigma_1 + \sigma_2) = 1 - w_1$$
- When these weights are chosen then  $\sigma_p = 0$ .

### EFFICIENT AND MINIMUM VARIANCE FRONTIERS

- On the following figure we introduce the
1. **Minimum variance frontier** and the
  2. **Efficient frontier** on the **portfolio opportunity set**.

### EFFICIENT AND MINIMUM VARIANCE FRONTIERS



### EFFICIENT AND MINIMUM VARIANCE FRONTIERS

- Notice that on the **minimum variance frontier** for each standard deviation there are two alternative expected returns (a high and a low expected return).
- The **efficient frontier** contains only the higher expected return risky portfolio.
- Thus, the solution of Step 1 is on the efficient frontier.

### OPTIMAL RISKY PORTFOLIO

- How to choose the optimal risky portfolio from the efficient frontier?
- First, define the **Sharp-ratio** as:  

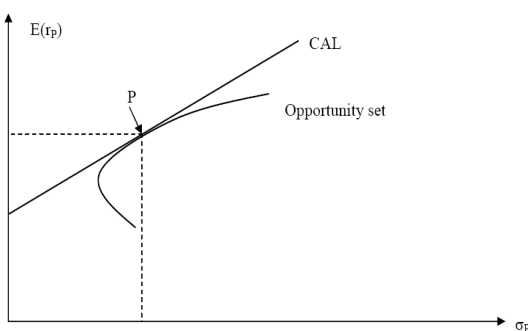
$$\text{Sharp-ratio} = [E(r_P) - r_f] / \sigma_P$$
- Then, we choose the highest **Sharp-ratio** portfolio  $P$  of the two risky assets, which we call the **optimal risky portfolio**.

### OPTIMAL RISKY PORTFOLIO

#### Optimal risky portfolio, $P$ :

- Consider that the risk-free rate is  $r_f$ .
- Graph the portfolio opportunity set and find the risky portfolio with highest **Sharp-ratio**.

### OPTIMAL RISKY PORTFOLIO



### OPTIMAL RISKY PORTFOLIO

- $P$  is the portfolio with the highest **Sharp-ratio** of the portfolio that contains the two risky assets.

## OPTIMAL RISKY PORTFOLIO

### Solution

- In the case of two risky assets, the solution for the weights of the optimal risky portfolio,  $P$  is the following:

$$w_1 = \frac{\{ [E(r_1) - r_f] \sigma_2^2 - [E(r_2) - r_f] \text{Cov}(r_1, r_2) \}}{\{ [E(r_1) - r_f] \sigma_2^2 + [E(r_2) - r_f] \sigma_1^2 - [E(r_1) - r_f + E(r_2) - r_f] \text{Cov}(r_1, r_2) \}}$$

$$w_2 = 1 - w_1$$

## OPTIMAL RISKY PORTFOLIO

### Generalization to many risky assets

- The previous slides showed the formulas for the 2-risky asset case.
- To optimize a portfolio of many risky assets, one may use numerical search methods to find the optimal values of  $w_i$  with  $i = 1, \dots, N$  such that the Sharp ratio of the portfolio is maximized.
- (See the practice of portfolio theory.)

## STEP 2: OPTIMAL COMPLETE PORTFOLIO

## OPTIMAL COMPLETE PORTFOLIO

- How to combine optimally the optimal risky portfolio with the risk-free asset?
- We shall combine the risk-free asset with portfolio  $P$  in order to determine the complete portfolio with **highest utility**.
- This way, we determine the **optimal complete portfolio**,  $C$ .
- The choice of the optimal complete portfolio is called **capital allocation decision**.

## THE RISK FREE ASSET

- Before entering into the details of the capital allocation decision, we review the definition and the properties of the risk-free asset.
- It is a common practice to view Treasury bills (T-bills) as the risk free asset.
- The reason is that only the government has the power to tax and control the money supply and so issue default-free bonds.

## THE RISK FREE ASSET

- However, even the default-free guarantee by itself is not sufficient to make the bonds risk-free in real terms.
- The only risk-free asset **in real terms** would be a perfectly **price-indexed bond**
- Price-indexed means that the bond is indexed against the inflation.

### THE RISK FREE ASSET

- Moreover, the price-indexed bond offers a guaranteed real rate to the investor **only if the maturity of the bond is equal to the investor's desired holding period.**
- Therefore, risk-free asset in real terms does not exist in practice.
- It only exists in **nominal terms.**

### THE RISK FREE ASSET

- In practice, most investors use **money market instruments** as a risk-free asset.
- These assets are virtually free of any interest rate risk because of their short maturities and because they are safe in terms of default or credit risk.

### THE RISK FREE ASSET

- **Money market funds** for most part contain three assets:
  - **Treasury bills** – issued by the government
  - **Bank certificates of deposit (CD)** – issued by banks
  - **Commercial papers (CP)** – issued by well-know companies

### Capital allocation decision

- **Capital allocation decision:** the choice of the proportion of the risk-free security and the optimal risky portfolio.
- The investor wants to choose the proportion of the **optimal risky portfolio**,  $y$  and that of the **risk-free asset**,  $1-y$ .

### Capital allocation decision

- Denote  $f$  the risk-free asset,  $P$  the risky portfolio and  $C$  the complete portfolio of the risky and the risk-free assets.
- Let  $r_f$  denote the rate of return of the risk-free asset, let  $E(r_P)$  denote the expected return of the risky portfolio and let  $E(r_C)$  be the expected return of the complete portfolio.
- Moreover, let  $\sigma_P$  denote the standard deviation of the risky portfolio and let  $\sigma_C$  be that of the complete portfolio.

### Capital allocation decision

- Then, the expected return and the risk of the complete portfolio,  $C$  can be written as follows:

$$E(r_C) = y E(r_P) + (1-y) r_f \quad (1)$$

and

$$\sigma_C^2 = y^2 \sigma_P^2 \quad (2)$$

### Capital allocation decision

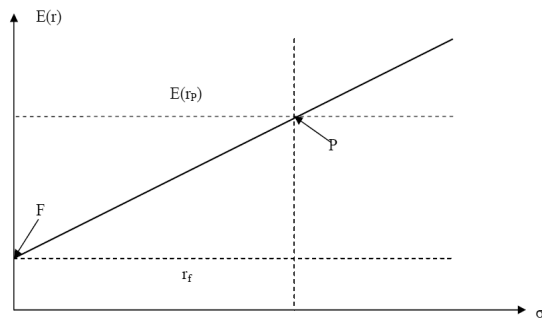
- Substituting (2) into (1) and rearranging the equation we get the expression of  $E(r_C)$  as a function of  $\sigma_C$ :  

$$E(r_C) = r_f + \sigma_C [E(r_P) - r_f] / \sigma_P \quad (3)$$
- This equation forms the **capital allocation line**.

### Investment opportunity set or capital allocation line (CAL)

Graph equation (3) in the following figure:

### Capital allocation line (CAL) $E(r_C)$ as a function of $\sigma_C$



### Capital allocation line (CAL)

- The slope of the CAL is the proportion of the risk premium and the standard deviation of the risky portfolio:  

$$[E(r_P) - r_f] / \sigma_P$$
- Notice that the slope of the CAL is the **Sharp-ratio**.

### Capital allocation line (CAL)

- The CAL pictures all possible complete portfolios between  $F$  and  $P$ .
  - When  $y = 0$  then we are in  $F$ , the risk-free asset.
  - When  $y = 1$  then we obtain  $P$ , the risky portfolio.
  - When  $0 < y < 1$  then we are between  $F$  and  $P$  in the line.

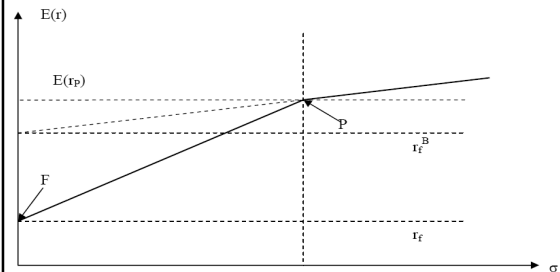
### Capital allocation line (CAL)

- What about points to the right of portfolio  $P$ ?
- These portfolios can be obtained by borrowing an additional amount from the risk-free asset. In case of borrowing,  $y > 1$ , and the complete portfolio is to the right of  $P$  on the figure.

### Capital allocation line (CAL)

- However, non-government institutions cannot borrow at the risk-free rate. Investors can borrow on interest rates higher than the risk-free rate in order to buy additional risky assets.
- We can picture this situation on the following graph:

### Capital allocation line (CAL)



- On this figure,  $r_f^B$  is the borrowing rate of the credit. To the right from  $P$  the slope of the CAL will be  $[E(r_P) - r_f^B] / \sigma_P$ .

### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- Individual investors' differences in risk aversion imply that, given an identical CAL set, different investors will choose different positions on the figure.
- In particular, the more risk averse investors will tend to hold less risky asset and more risk-free asset.

### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- Thus, the optimal choice will depend on the utility function of the risk averse investor:

$$U = E(r) - A \sigma^2$$

where  $A > 0$  is an index of the investor's risk aversion.

### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- As the investor wants to maximize its utility obtained from the complete portfolio of the risky and risk-free assets, we have the following maximization problem to be solved:

$$\max_y U(y) = E(r_C) - A \sigma_C^2$$

### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- Substituting

$$E(r_C) = y E(r_P) + (1-y) r_f \quad (1)$$

and

$$\sigma_C^2 = y^2 \sigma_P^2 \quad (2)$$

into the optimization problem we obtain:

$$\max_y U(y) = \max_y E(r_C) - A \sigma_C^2 = \max_y r_f + y[E(r_P) - r_f] - A y^2 \sigma_P^2$$

### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- We can solve this problem by taking derivative with respect to  $y$  and equal it to zero.
- The solution of the problem is the following:

$$y^* = [E(r_P) - r_f] / 2A \sigma_P^2$$

### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

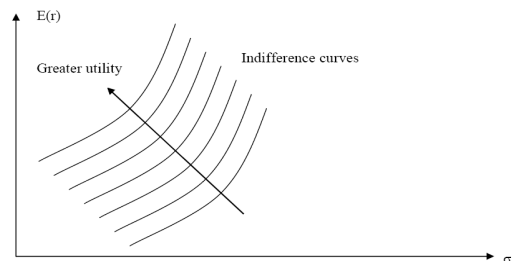
- Thus, the result obtained is intuitive:
  - (1) Higher risk aversion,  $A$  implies lower investment in the risky asset,
  - (2) Higher risk premium,  $[E(r_P) - r_f]$  implies higher investment in the risky portfolio, and
  - (3) Higher risk of  $P$ ,  $\sigma_P$  implies lower investment in the risky portfolio.

### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- A graphical way of presenting this decision problem is to use indifference curve analysis.
- Recall that the indifference curve is a graph in the expected return – standard deviation plane of all points that result in equal level of utility.
- The curve displays the investor's required trade-off between expected return and standard deviation (risk).

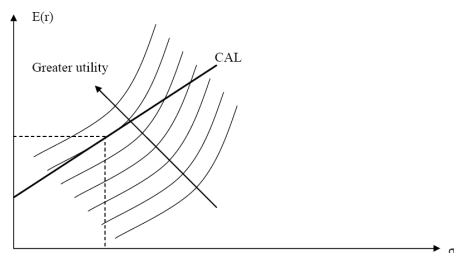
### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- First, picture different indifference curves corresponding to higher utility values:



### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- Second, include the CAL investment opportunity line into the figure:

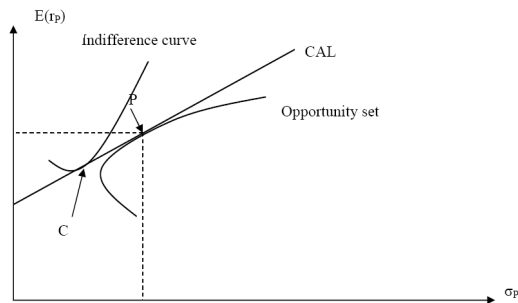


### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- The investor seeks the position with the highest feasible level of utility, represented by the highest possible indifference curve that touches the investment opportunity set (CAL).
- This is the indifference curve tangent to the CAL:



### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL



### OPTIMAL COMPLETE PORTFOLIOS FROM THE CAL

- The figure shows that the optimal complete portfolio is determined by the point where the slope of the indifference curve is equal to the CAL.

### PASSIVE STRATEGIES: THE CAPITAL MARKET LINE (CML)

- In the previous section, we used the optimal risky portfolio  $P$  in order to determine the CAL.
- The choice of  $P$  would require some analysis of the capital market.

### CAPITAL MARKET LINE (CML)

- One possibility would be to avoid doing any analysis and follow the so-called **passive strategy**.
- This means that  $P$  would represent a broad index of risky assets, for example the S&P500 or IBEX-35 stock index.
- In this case,  $P$  is chosen without any capital market analysis and the resulting CAL is called the **capital market line (CML)**.

### CAPITAL MARKET LINE (CML)

- Why would an investor follow the passive strategy of asset allocation?
  - (1) **It is cheap:** The alternative active strategy is **not free**. For the capital market analysis the investor has **fees** and other **costs**.
  - (2) **Free rider benefit:** In a competitive capital market, a well-diversified portfolio of common stocks will be a reasonably **fair buy**, and the passive strategy may not be inferior to that of the average active investor.
- That is by the passive strategy we are free riding on active knowledgeable investors who make stock prices a fair buy.

### NOTE: THE SEPARATION PROPERTY

- The determination of the optimal risky portfolio  $P$  is independent of the preferences of the investors.
- Therefore, the portfolio manager will offer the same  $P$  to all clients regardless of their degree of risk aversion.
- Thus, the solution of step (1) and (2) can be separated completely.
- This is called the **separation property**.

#### NOTE: THE SEPARATION PROPERTY

- The separation property makes professional management more efficient and less costly.
  - Step 1: the determination of the optimal risky portfolio,  $P$  is purely technical.
  - Step 2: the determination of the optimal complete portfolio,  $C$  depends on the client's preferences.

#### NOTE: WHAT TO DO IF WE ARE NOT SURE ABOUT THE VALUE OF $A$ ?

- One possibility is to solve the complete portfolio optimization problem for alternative values of  $A$ .
- Then, we review the obtained alternative optimal portfolio corresponding to different  $A$  values.
- Finally, the client may choose according to his/her preferences among the offered solutions.

## Forecasting security prices

### STRUCTURE OF CLASS

1. Motivation
2. Definition of forecast
3. Procedure of forecasting
4. Econometric models used for forecasting
5. Evaluation of forecast precision
6. Additional remarks (out-of-sample and in-sample forecasting, one-step-ahead and multi-step-ahead forecasts, stationarity conditions)

### MOTIVATION

- Investors and financial analysts are frequently interested in forecasting prices of financial assets.
- Bank analysts frequently write reports to their clients about expected prices of financial products.

### Definition of forecast

### DEFINITION OF FORECAST

- In this class, we are interested in the **direction** of the price change.
- We model the log return  $y_t$  on the investment at time  $t$ :
$$y_t = \ln(p_t/p_{t-1})$$
- We consider  $t = 1, 2, 3, \dots, T$  time periods.

### DEFINITION OF FORECAST

- The objective is to forecast the return for period  $t$ ,  $y_t$  given all previous information.
- When we compute our forecast, we are in the beginning of period  $t$ .
- In that moment,
  - We know  $(y_1, \dots, y_{t-1})$  but
  - We do not know  $y_t$

### DEFINITION OF FORECAST

- What does the word “**forecast**” mean for us?
- The **forecast** of the variable  $y_t$  for the period  $t$  is defined as the **mean** of  $y_t$  given all past information observed until the end of period  $t-1$  ( $t-1$  included).
- We denote the forecast by  $\mu_t$

### DEFINITION OF FORECAST

- Mathematically, the forecast can be formalized as

$$\text{Forecast of } y_t = \mu_t = E[y_t | F_{t-1}]$$

where  $F_{t-1}$  denotes all past information observed until time  $t-1$  ( $t-1$  included).

### DEFINITION OF FORECAST

- $F_t$  denotes **all past information observed** until the end of period  $t$ .
- For example, suppose that all past information used to make the forecast are past values of returns that is
$$F_{t-1} = (y_1, \dots, y_{t-1})$$
- Then the forecast can be written as
$$E[y_t | y_1, y_2, \dots, y_{t-1}]$$

### DEFINITION OF FORECAST

- In general, a forecast can be done using more information than only the past price data.
- We may use additional explanatory variables denoted  $x_t$  to estimate the future return if we think that the explanatory variables contain important information on future price movements.

### DEFINITION OF FORECAST

- If we use past values of the additional explanatory variables to forecast than the **information set** is:
$$F_t = (y_1, \dots, y_{t-1}, x_1, \dots, x_{t-1})$$
- In this case, the forecast of return  $y_t$  can be written as
$$E[y_t | y_1, \dots, y_{t-1}, x_1, \dots, x_{t-1}]$$

### Procedure of forecasting

### Procedure of forecasting

- (1) **Collect past data** on the financial prices to be forecasted ( $y_1, \dots, y_{t-1}$ ) and collect past data on the additional explanatory variables of interest ( $x_1, \dots, x_{t-1}$ ).
- (2) **Select an econometric model** for  $y_t$ .
- (3) **Estimate** the parameters of the selected econometric model.

### Procedure of forecasting

- (4) **Compute**  $\mu_t$  using past values of returns, explanatory variables and the parameters estimates of the econometric model.
- (5) **Evaluate the forecast performance** comparing the realized value of  $y_t$  and the forecasted value  $\mu_t$ .

### Econometric models used for forecasting

#### Econometric models used for forecasting

- There are alternative econometric models that can be used for forecasting purposes.
- For each model, we show its specification and the computation of the conditional mean of future returns, i.e. the forecast formula.

#### Econometric models used for forecasting

- We are going to present very general models that may include several lags of the variables.
- However, including many variables have two opposite effects on forecast precision:

#### Econometric models used for forecasting

- (1) More variables mean more past information used to forecast.  
This is a POSITIVE effect.

### Econometric models used for forecasting

(2) More variables mean more parameters to be estimated.

- This reduces the precision of the statistical estimation of the model. (This means that the estimated parameter value may be far from the true value.)

This is NEGATIVE effect.

- In practice, we need to find the correct balance between these two effects.

### Econometric models used for forecasting

- We review various econometric models of  $\mu_t$ .
- In each model, we suppose that the volatility of the security  $\sigma$  is constant and that the return given all past information is normally distributed:

$$y_t | F_t \sim N(\mu_t, \sigma^2)$$

### Econometric models used for forecasting

- We suggest the following models for the conditional mean:

1. AR( $p$ )
  2. ARMA( $p, q$ )
  3. AR( $p$ )-X( $k$ )
  4. ARMA( $p, q$ )-X( $k$ )
- AR( $p$ ) = autoregressive of order  $p$
  - ARMA( $p, q$ ) = autoregressive (AR) of order  $p$  and moving average (MA) of order  $q$
  - X( $k$ ) =  $k$  lags of the exogenous variable

#### 1. AR( $p$ ) model

- **AR( $p$ ) model:**

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + u_t$$

where  $u_t$  is the i.i.d  $N(0, \sigma^2)$  error term.

- The parameters are  $c, \phi_i$  with  $i=1, \dots, p$  and  $\sigma$ .

#### 1. AR( $p$ ) model

- The one-step-ahead forecast formula is

$$\mu_t = c + \sum_{i=1}^p \phi_i y_{t-i}$$

#### 2. ARMA( $p, q$ ) model

- **ARMA( $p, q$ ) model:**

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j u_{t-j} + u_t$$

where  $u_t$  is the i.i.d  $N(0, \sigma^2)$  error term.

- The parameters are  $c, \phi_i$  with  $i=1, \dots, p, \theta_j$  with  $j=1, \dots, q$  and  $\sigma$ .

## 2. ARMA( $p,q$ ) model

- The one-step-ahead forecast formula is

$$\mu_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j u_{t-j}$$

## 3. AR( $p$ )-X( $k$ ) model

- **AR( $p$ )-X( $k$ ) model:**

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{l=1}^k m_l X_{t-l} + u_t$$

where  $u_t$  is the i.i.d  $N(0, \sigma^2)$  error term.

The parameters are  $c, \phi_i$  with  $i=1, \dots, p$ ,  $m_l$  with  $l=1, \dots, k$  and  $\sigma$ .

## 3. AR( $p$ )-X( $k$ ) model

- The one-step-ahead forecast formula is

$$\mu_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{l=1}^k m_l X_{t-l}$$

## 4. ARMA( $p,q$ )-X( $k$ ) model

- **ARMA( $p,q$ )-X( $k$ ) model:**

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j u_{t-j} + \sum_{l=1}^k m_l X_{t-l} + u_t$$

where  $u_t$  is the i.i.d  $N(0, \sigma^2)$  error term.

- The parameters are  $c, \phi_i$  with  $i=1, \dots, p$ ,  $\theta_j$  with  $j=1, \dots, q$ ,  $m_l$  with  $l=1, \dots, k$  and  $\sigma$ .

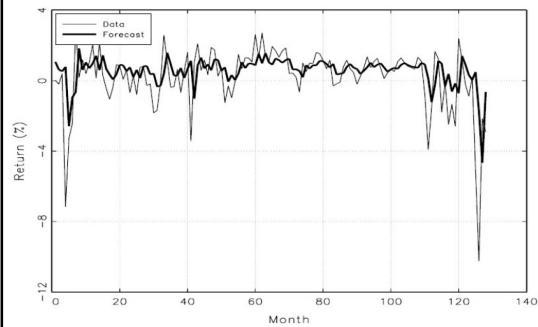
## 4. ARMA( $p,q$ )-X( $k$ ) model

- The one-step-ahead forecast formula is

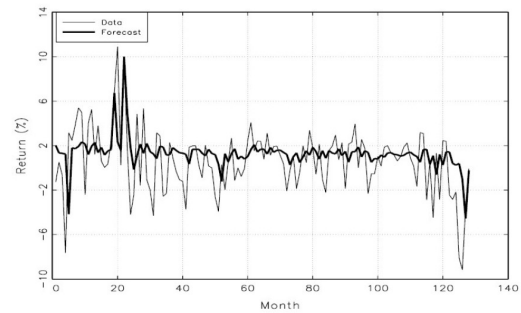
$$\mu_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j u_{t-j} + \sum_{l=1}^k m_l X_{t-l}$$

## EXAMPLES OF SOME FORECASTS OF HEDGE FUND INDEX LOG RETURNS

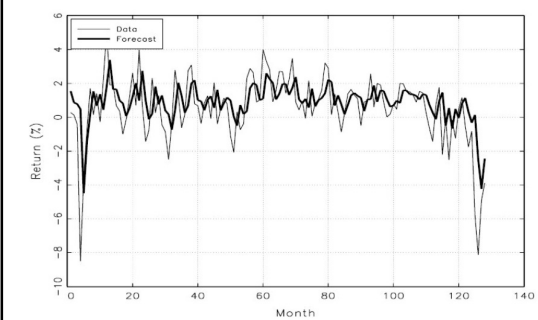
### Some examples of forecasts



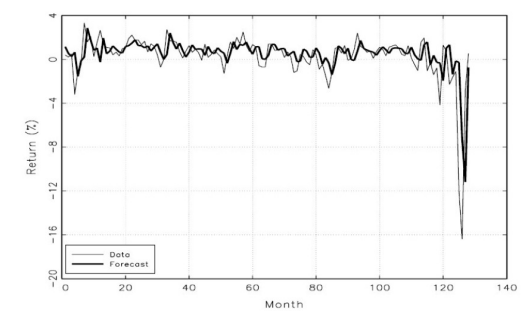
### Some examples of forecasts



### Some examples of forecasts



### Some examples of forecasts



### EVALUATION OF FORECAST PRECISION

#### Evaluation of forecast precision

- Several alternative measures of forecast precision exist.
- These measures compare the distance of the true time series  $y_t$  and the forecasted time series  $\mu_t$ .



### Evaluation of forecast precision

- We present three alternative forecast performance measures:
  - (1) **Mean absolute error (MAE)**
  - (2) **Mean square error (MSE)**
  - (3) **Root mean square error (RMSE)**

### Evaluation of forecast precision

- The MAE, MSE and RMSE measures are formalized as follows:

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |y_t - \mu_t|$$

$$\text{MSE} = \frac{1}{T} \sum_{t=1}^T (y_t - \mu_t)^2$$

$$\text{RMSE} = \sqrt{\text{MSE}}$$

### Evaluation of forecast precision

- An advantage of the MAE and RMSE measures is that the scale of both measures is the same as the scale of the variable of interest that is forecasted,  $y_t$ .
- The disadvantage of the MSE measure is that the scale of the MSE is different to the scale of  $y_t$ .

### Additional remarks

### Out-of-sample and in-sample forecasting

- There are two ways to perform forecasting:
  1. **In-sample forecast**
  2. **Out-of-sample forecast**

### In-sample forecast

- Suppose that we observe  $t = 1, \dots, T$  periods of returns.
- In the **in-sample forecast**, we estimate an econometric model using all data covering the period  $t = 1, \dots, T$ .
- Then, we compute  $\mu_t$  inside the  $t = 1, \dots, T$  period.
- This forecast procedure is **not realistic** as **we use future information** to estimate to parameters of the econometric model.

### Out-of-sample forecast

- Suppose that we observe  $t = 1, \dots, T$ .
- In the **out-of-sample forecast**, we estimate an econometric model using data for the period  $t = 1, \dots, T$ .
- Then, we forecast the return for next period  $t = T+1$ .
- This forecast procedure is **more realistic** as here **we use past information** to estimate the parameters of the econometric model.

### One-step-ahead forecasts

- In the previous slides, we presented formulas for the one-step-ahead forecasts:
- $$E[y_{t+1}|y_1, y_2, \dots, y_t]$$
- In the one-step-ahead forecast, we are only interested in the forecast of the next period  $t+1$  and we are not interested in forecasting further periods  $t+2, t+3, \dots$

### Multi-step-ahead forecasts

- However, in some situations it may be interesting to forecast for further periods.
- For example, we may need estimates of:
 
$$E[y_{t+2}|y_1, y_2, \dots, y_t]$$

$$E[y_{t+3}|y_1, y_2, \dots, y_t]$$
- These forecasts are called multi-step-ahead forecasts.
- In the example, these are two-step-ahead and three-step-ahead forecasts.

### Multi-step-ahead forecasts

- Imagine that we are in the beginning of period  $t$  and we want to do a multi-step-ahead forecast of future log returns, i.e. we know  $(y_1, \dots, y_{t-1})$  but do not know  $(y_t, \dots, y_T)$ .
- In the followings, we show the formula for the AR(1) and ARMA(1,1) models.
- We show that this formula depends on the one-step-ahead forecast.

### Multi-step-ahead forecast for AR(1) and ARMA(1,1)

- The multi-step-ahead forecasts for the AR(1) and ARMA(1,1) models can be computed recursively as:

$$E(y_{t+1}|F_{t-1}) = c + \phi E(y_t|F_{t-1})$$

- where the expectation on the right hand side is different for each model:

$$\text{AR}(1): E(y_t|F_{t-1}) = c + \phi y_{t-1}$$

$$\text{ARMA}(1,1): E(y_t|F_{t-1}) = c + \phi y_{t-1} + \theta u_{t-1}$$

### Stationarity of AR(1) and ARMA(1,1) models

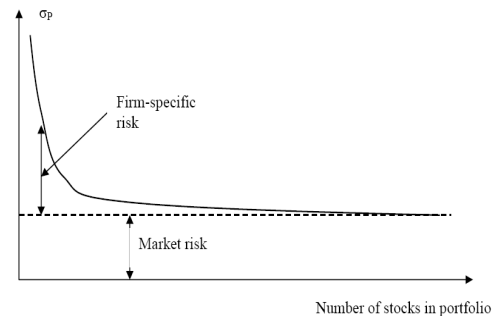
- The estimated AR(1) and ARMA(1,1) models have to be covariance stationary in order to explain properly the data.
- Stationarity can be verified by looking at the estimated autoregressive coefficient,  $\phi$ .
- Both for the AR(1) and ARMA(1,1) models, the stationarity condition is:  $|\phi| < 1$ .
- For ARMA(p,q), the roots of the polynomial

$$1 - \sum_{i=1}^p \phi_i z^i = 0$$

are outside the unit circle

## Factor models

### MOTIVATION 1 – MARKET RISK



### MOTIVATION 1 – MARKET RISK

- When common sources of risk affect all firms, even extensive diversification cannot eliminate risk.
- The risk that remains even after extensive diversification is called **market risk**.
- The factor models can be used to **estimate** the market risk and the firm-specific risk components of security returns.

### MOTIVATION 2 – MARKOWITZ

- The Markowitz portfolio selection model uses the following inputs to form optimal portfolios:
  - (1) expected return of each security
  - (2) variance-covariance matrix of security returns
- These inputs the analyst should estimate from empirical data.
- In the followings, we show how many parameters are to be estimated in the Markowitz model.

### MOTIVATION 2 – MARKOWITZ

- We need to estimate:
  - $N$  expected returns +
  - $N$  standard deviations +
  - $(N^2 - N)/2$  correlation coefficients =
- TOTAL =  $2N + (N^2 - N)/2$  parameters

### MOTIVATION 2 – MARKOWITZ

- If  $n = 1,600$  (roughly the number of stocks at New York Stock Exchange, NYSE) then TOTAL = 1.3 million parameters to be estimated.
- To estimate 1.3million different parameters is impossible from statistical point of view because the number of data observed is much less than this value.

### MOTIVATION 2 – MARKOWITZ

- During the past decades, alternative models have appeared in finance, which can estimate the inputs for the Markowitz framework by relatively low number of parameters.
- A class of these models is called **factor models**.

### MOTIVATION 2 – expected return

- The factor models specify the **expected excess return** or **expected risk premium** of risky securities.
- Excess return =  
asset return – risk-free rate
- Expected excess return =  $E(r_i) - r_f$
- A factor model provides an estimate of the expected excess return  $\mu$  of a risky security.

### MOTIVATION 2 – expected return

- In the factor model specifications, the returns of individual securities are driven by one or more **common factors**.
- These factors are common to all securities but their influence on each asset can be different.
- The specific impact of a factor on individual security return is measured by the '**beta**' coefficient.

### MOTIVATION 2 – covariance

- The inclusion of common factors makes possible the estimation of the covariance matrix of several assets' returns.
- Moreover, the number of parameters to be estimated in  $\Sigma$  are low, which makes feasible these models in practice.

### MOTIVATION 2 – MARKOWITZ

- By a factor model we can estimate the  $\mu$  and  $\Sigma$  of **returns**, which can be used as inputs to the portfolio optimization problem.
- This fact motivates the study of factor models.

### FACTOR MODELS

- We consider four factor models:
  - (1) **Capital asset pricing model (CAPM)**
  - (2) **Index model with one market factor (single-index model)**
  - (3) **Fama-French 3-factor model**
  - (4) **Arbitrage pricing theory (APT)**

## FACTOR MODELS

- For each model, we cover MOTIVATION 2 first.
- Then, in a general setup we present the solution for MOTIVATION 1.

## (1) CAPITAL ASSET PRICING MODEL

### Capital Asset Pricing Model (CAPM)

- The CAPM is a central model of modern financial economics.
- The CAPM derives that the expected return of a security is driven by a common 'market' risk premium.
- It provides a benchmark rate of expected return for evaluating possible investments of given risk.

### The CAMP 'beta'

- A central concept of CAPM is the 'beta' coefficient, which is security specific.
- **Definition (beta):** The beta coefficient measures the extent to which returns on the stock and the market move together:
$$\beta_i = \text{Cov}(r_i, r_M) / \sigma_M^2$$
- Beta can be seen as an **alternative measure of financial risk** because it gives the **sensitivity** of the single asset to the market return.

### CAPM central result

- In the CAPM, the risk premium on individual *assets* is:

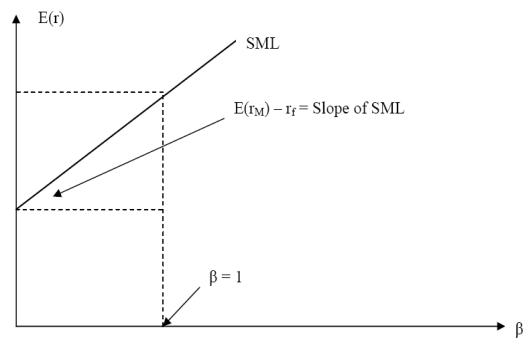
$$E(r_i) = r_f + \beta_i [E(r_M) - r_f]$$

- This is the CAPM equation known by practitioners.

### Security market line - SML

- The expected return – beta relationship of CAPM can be plotted graphically as the **security market line** (SML):

### Security market line - SML



### Security market line - SML

- The SML provides a benchmark for the evaluation of investment performance:
- Given the risk of an investment, as measured by its beta, the SML provides the required rate of return from that investment to compensate investors for risk, as well as the time value of money.

### Security market line - SML

- Because the security market line is the graphic representation of the expected return – beta relationship, “fairly priced” assets plot exactly on the SML.
- The difference between the fair and the actually expected rates of return is called the stock’s alpha, denoted  $\alpha_i$ .

### Beta of a portfolio

- If the expected return – beta relationship holds for any individual asset, it must hold for any combination of assets.
- The portfolio beta is given by the weighted average of individual betas  $\beta_i$ :

$$\beta_P = \sum_{i=1}^N w_i \beta_i$$

where  $w_i$  denotes the weight of the  $i$ -th asset.

### Aggressive/defensive stocks

- Betas in absolute value greater than 1 are considered **aggressive** because high-beta stocks entails above-average sensitivity to market swings.
- Betas in absolute value lower than 1 can be described as **defensive** investments because low-beta stocks entails below-average sensitivity to market swings.

### (2) Single-index model

### Single-index model

- The **single-index model** is formulated as follows:  

$$r_i - r_{ft} = \alpha_i + \beta_i (r_{Mt} - r_{ft}) + e_{it}$$
- The single-index model uses the market risk premium,  $(r_M - r_f)$  as a common factor, which has different impact on each security measured by the  $\beta_i$ .
- The  $e_{it}$  is a zero mean error term with variance  $(\sigma_{e_{it}})^2$
- $\alpha_i$  is a security specific constant term.
- Notice that this model is very similar to the CAPM.

### Estimating the index model

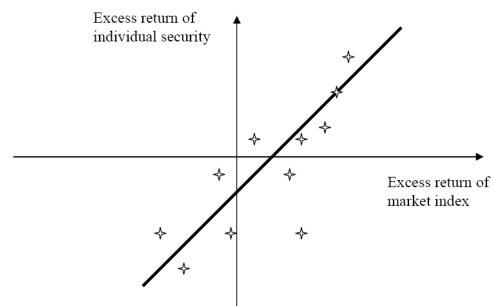
- In a time series setup, we can estimate this equation for asset  $i$  by as a **linear regression model**:  

$$r_{it} - r_{ft} = \alpha_i + \beta_i (r_{Mt} - r_{ft}) + e_{it}$$
- We estimate this regression for each asset  $i=1, \dots, N$  to obtain the parameter estimates.

### Estimating the index model

- If we plot firm-specific excess returns as a function of market excess returns using the regression estimates for empirical data we get the following **security characteristic line (SCL)**:

### Estimating the index model



### Expected return of assets in CAPM and single-index models

- Given the estimation results of the linear regression, one can compute the expected return of a security as follows:  

$$E[r_{it}] = \alpha_i + \beta_i (E[r_{Mt}] - E[r_{ft}]) + E[r_{ft}]$$
for each  $i=1, \dots, N$ .

### Covariance matrix of returns in the CAPM and the single-index model

- We can express the covariance matrix between asset returns as follows:
- Let  $\beta$  be an  $2 \times N$  matrix of betas of all assets in the portfolio:  

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \beta_1 & \beta_2 & \dots & \beta_N \end{bmatrix}$$
- Let  $\Sigma$  denote the  $2 \times 2$  variance-covariance matrix of  $r_{ft}$  and  $(r_{Mt} - r_{ft})$ .

Covariance matrix of returns in the CAPM and the single-index model

- Then, the covariance matrix of asset returns is given by

$$\beta' \Sigma \beta + I_N (\sigma_e)^2$$

- where  $I_N$  denotes the  $N \times N$  identity matrix,
- $(\sigma_e)^2$  denotes the  $N \times 1$  vector of variances of the error terms  $e_{it}$

### (3) Fama-French 3-factor model

#### Fama-French 3-factor model

- The Fama-French 3-factor model is formulated as follows:

$$r_i - r_f = \alpha_i + \beta_{i1} F_1 + \beta_{i2} F_2 + \beta_{i3} F_3 + e_i$$

where the 3 factors are

- $F_1 = (r_M - r_f)$  = market risk premium
- $F_2$  = SMB, 'small minus big'
- $F_3$  = HML, 'high minus low'

- The model includes a zero mean error term  $e_i$  and a security specific constant  $\alpha_i$ .

#### Fama-French factor data

- Fama-French factors can be downloaded from the following website:  
[http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)

#### Fama-French 3-factor model

- The three factors include all NYSE, AMEX, and NASDAQ firms for which there is market and book equity data is available.

#### Construction of Fama-French factors

- The Fama/French factors are constructed using the 6 value-weight portfolios formed on size and book-to-market:
- 2 groups according to size:
  - small
  - big
- 3 groups according to book-to-market of equity:
  - value (high book/market)
  - neutral
  - growth (low book/market)



### Construction of $(r_m - r_f)$

- (1) The excess return on the market, is the value-weight return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate.

### Construction of SMB

- (2) SMB is the average return on the three small portfolios minus the average return on the three big portfolios,  

$$OSMB = \frac{1}{3} (\text{Small Value} + \text{Small Neutral} + \text{Small Growth}) - \frac{1}{3} (\text{Big Value} + \text{Big Neutral} + \text{Big Growth}).$$

### Construction of HML

- (3) HML is the average return on the two value portfolios minus the average return on the two growth portfolios,  

$$OHML = \frac{1}{2} (\text{Small Value} + \text{Big Value}) - \frac{1}{2} (\text{Small Growth} + \text{Big Growth}).$$

### Estimation of the Fama-French model

- In a time series setup, the alpha and betas of the model for security  $i$  can be estimated by **linear regression** on the model:

$$r_{it} - r_{ft} = \alpha_i + \beta_{i1} F_{1t} + \beta_{i2} F_{2t} + \beta_{i3} F_{3t} + e_{it}$$

- We estimate this regression for each asset  $i=1, \dots, N$  to get the corresponding parameters.

### Expected return of assets in the Fama-French model

- Given the estimation results of the linear regression, one can compute the expected return of a security as follows:

$$E[r_{it}] = \alpha_i + \beta_{i1} E[F_{1t}] + \beta_{i2} E[F_{2t}] + \beta_{i3} E[F_{3t}] + E[e_{it}]$$

for each  $i=1, \dots, N$ .

### Covariance matrix of returns in the Fama-French model

- We can express the covariance matrix between asset returns as follows:
- Let  $\beta$  be an  $4 \times N$  matrix of betas of all assets in the portfolio:

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \beta_{11} & \beta_{21} & \dots & \beta_{N1} \\ \beta_{12} & \beta_{22} & \dots & \beta_{N2} \\ \beta_{13} & \beta_{23} & \dots & \beta_{N3} \end{pmatrix}$$

Covariance matrix of returns in the Fama-French model

- Let  $\Sigma$  denote the 4 x 4 variance-covariance matrix of  $r_{it}$ ,  $F_{1t}$ ,  $F_{2t}$  and  $F_{3t}$
- Then, the covariance matrix of asset returns is given by  

$$\beta' \Sigma \beta + I_N (\sigma_e)^2$$
- where  $I_N$  denotes the  $N \times N$  identity matrix
- $(\sigma_e)^2$  denotes the  $N \times 1$  vector of variances of the error terms  $e_{it}$

## (4) ARBITRAGE PRICING THEORY

### MOTIVATION

- Theoretically and empirically, one of the most troubling problems of CAPM for academics and managers has been that the CAPM's single source of risk is the market.
- As a consequence, both academics and practitioners have analyzed the influence of adding additional factors into the CAPM equation.

### MOTIVATION

- The most famous multi-factor model, the **arbitrage pricing theory (APT)**, was developed by Ross (1976).

### Arbitrage Pricing Theory (APT)

- The risk premium of a risky asset in APT can be written as

$$r_i - r_f = \alpha_i + \beta_{i1} F_1 + \dots + \beta_{im} F_m + e_i$$

where  $F_j$ ,  $j = 1, \dots, m$  denotes the factor, that is, the systematic component of the return, and  $e_i$  corresponds to the firm-specific component of the return.

### Factors of APT

- The APT does not say what the factors are.
- For example, they could be an oil price factor or an interest rate factor.
- The return on the market portfolio might serve as one factor, as in the CAPM, but it might not as well.

### Estimation of APT

- In a time series setup, the APT equation can be estimated for asset  $i$  by linear regression model:

$$r_{it} - r_{ft} = \alpha_i + \beta_{i1}F_{1t} + \dots + \beta_{im}F_{mt} + e_{it}$$

- We estimate this regression for each asset  $i=1, \dots, N$  to get the parameter estimates.

### Expected return of assets in APT

- Given the estimation results of the linear regression, one can compute the expected return of a security as follows:

$$E[r_{it}] = \alpha_i + \beta_{i1}E[F_{1t}] + \dots + \beta_{im}E[F_{mt}] + E[r_{ft}]$$

for each  $i=1, \dots, N$ .

### Covariance matrix of returns in the Fama-French model

- We can express the covariance matrix between asset returns as follows:
- Let  $\beta$  be an  $(m+1) \times N$  matrix of betas of all assets in the portfolio:

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \beta_{11} & \beta_{21} & \dots & \beta_{N1} \\ \dots & \dots & \dots & \dots \\ \beta_{1m} & \beta_{2m} & \dots & \beta_{Nm} \end{pmatrix}$$

### Covariance matrix of returns in APT

- Let  $\Sigma$  denote the  $(m+1) \times (m+1)$  variance-covariance matrix of  $r_{ft}, F_{1t}, \dots, F_{mt}$ .
- Then, the covariance matrix of asset returns is given by
- $\beta' \Sigma \beta + I_N (\sigma_e)^2$
- where  $I_N$  denotes the  $N \times N$  identity matrix
- $(\sigma_e)^2$  denotes the  $N \times 1$  vector of variances of the error terms  $e_{it}$

### CONCLUSION - FACTOR MODELS – number of parameters to be estimated

- We report the number of parameters to be estimated in general and for  $N=1600$ ,  $m=5$  ( $m$  APT factors)
  - CAPM:  $2N+1 = 3201$
  - Single-index:  $3N+1 = 4801$
  - Fama-French:  $5N+9 = 8009$
  - APT with  $m$  factors:  $(m+2) \times N + m^2 = 11225$
- Compare these with 1.3 million parameters of the general setup.

### ESTIMATION OF MARKET RISK

### ESTIMATION OF MARKET RISK

- In the following slides, the estimation procedure for the firm-specific and market-specific risk components are presented.
- We consider the following general formulation for the factor model, which includes all previous models:

$$r_{it} = r_{ft} + \alpha_i + \beta_i' F_t + e_{it}$$

- Notice that  $F_t$  is a vector of factors.

### ESTIMATION OF MARKET RISK

- The parameters of this model can be estimated by OLS for the next equation:

$$r_{it} - r_{ft} = \alpha_i + \beta_i' F_t + e_{it}$$

### ESTIMATION OF MARKET RISK

- The variance of the security can be decomposed as follows, by using the independence property of the error term:

$$\text{Var}[r_{it}] = \text{Var}[r_{ft} + \alpha_i + \beta_i' F_t] + \text{Var}[e_{it}]$$

### ESTIMATION OF MARKET RISK

- Therefore, the firm-specific and market-specific variance are given by:

$$\text{Var}[e_{it}] = \text{Firm-specific variance}$$

$$\text{Var}[r_{ft} + \alpha_i + \beta_i' F_t] = \text{Market-specific variance}$$

### ESTIMATION OF MARKET RISK

- The firm-specific component is given by the OLS estimation results.
- The market-specific component is given by

$$\text{Var}[r_{it}] - \text{Var}[e_{it}]$$

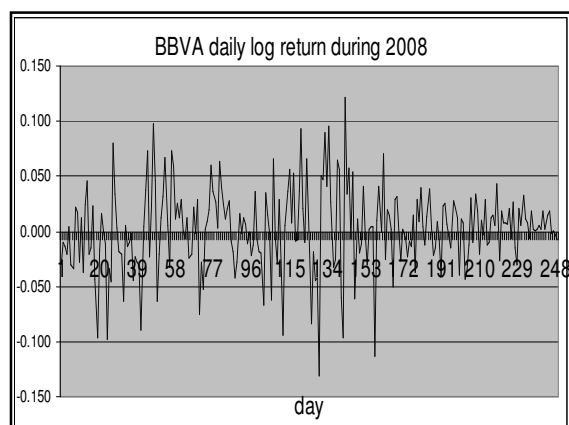
## Dynamic models of volatility

### MOTIVATION

- Both in the theory and practice of finance, volatility modelling is important.
- Volatility estimates are used for:
  - (1) **Risk management purposes** (for example to compute the value at risk, VAR of a portfolio).
  - (2) **Financial asset valuation purposes** (for example to determine the fair price of a stock option).
  - (3) **Portfolio construction purposes** (we need volatility estimates in order to construct the optimal risk-return portfolio).

### DEFINITION OF VOLATILITY

- In this section, volatility is defined as the standard deviation of asset log returns denoted by  $\sigma_t$
- The variance of log return is  $(\sigma_t)^2$
- Notice that volatility has sub-index  $t$ .
- There is empirical evidence showing the volatility is changing over time.
- On the next figure the BBVA stock log return evolution is presented during 2008:



### DYNAMIC VOLATILITY MODEL

- On the previous figure we can see 'high volatility periods' and 'low volatility periods'.
- This phenomenon can be modeled by a so-called dynamic volatility model.

### DYNAMIC VOLATILITY MODEL

#### General form of the model

- We are in the beginning of period  $t$ .
- We know  $(y_1, \dots, y_{t-1})$  but we do not know  $y_t$
- We model  $y_t$  as follows:

$$y_t | F_{t-1} \sim N(\mu_t, \sigma_t^2)$$

where  $F_{t-1} = (y_1, \dots, y_{t-1})$

### DYNAMIC VOLATILITY MODEL

- The term **mean equation** refers to a model of  $\mu_t$  – We modelled  $\mu_t$  in the forecasting topic.
- The term **volatility equation** refers to a model of  $\sigma_t^2$  – The volatility equation is the dynamic volatility model.
- In the followings, we present alternative models for  $\mu_t$  and  $\sigma_t^2$

### DYNAMIC VOLATILITY MODELS

- In particular, we are going to present various models of dynamic volatility:
  1. GARCH-type models
    - 1(a) ARCH,
    - 1(b) GARCH (special case: EWMA)
    - 1(c) EGARCH
  2. Stochastic volatility models

### 1(a) ARCH MODEL

#### GARCH-type volatility models ARCH

- The ARCH model has been developed by Engle (1982) and became very popular for the dynamic modelling of volatility.
- ARCH = autoregressive conditional heteroscedasticity
- There are different ARCH( $q$ ) models which differ in the number of lags included in the volatility equation.

#### GARCH-type volatility models ARCH

- The ARCH(1) model is formulated as:
$$y_t = \omega + \varepsilon_t$$
$$\varepsilon_t = \sigma_t u_t, \quad u_t \sim N(0,1)$$
$$(\sigma_t)^2 = \alpha_0 + \alpha_1 (\varepsilon_{t-1})^2$$
- where  $\alpha_i > 0$  for  $i=1,0$  to ensure the positive value of volatility and  $\omega$  is a real number.
- The  $(\sigma_t)^2$  denotes the variance of  $y_t$  and  $\sigma_t$  denotes the standard deviation of  $y_t$

#### GARCH-type volatility models ARCH

- The ARCH( $q$ ) model is formulated as follows:
$$y_t = \omega + \varepsilon_t$$
$$\varepsilon_t = \sigma_t u_t, \quad u_t \sim N(0,1)$$
$$(\sigma_t)^2 = \alpha_0 + \alpha_1 (\varepsilon_{t-1})^2 + \alpha_2 (\varepsilon_{t-2})^2 + \dots + \alpha_q (\varepsilon_{t-q})^2$$
- where  $\alpha_i > 0$  for  $i=1, \dots, q$  to ensure positive value of volatility and  $\omega$  is a real number.
- The  $(\sigma_t)^2$  denotes the variance of  $y_t$  and  $\sigma_t$  denotes the standard deviation of  $y_t$

### GARCH-type volatility models ARCH

#### Stationarity

- When a dynamic volatility model is estimated, it is important to check if the parameters estimates of the model determine a stationary or a non-stationary time series of volatility.
- The parameter estimates given by the econometric software **are correct only if** the model **is stationary**.
- On the following slide we give the conditions for stationarity.

### GARCH-type volatility models ARCH

#### Stationarity

- The ARCH( $q$ ) model is stationary if

$$\sum_{i=1}^q \alpha_i < 1$$

- The ARCH(1) model is stationary if  $\alpha_1 < 1$ .

## 1(b) GARCH MODEL

### GARCH-type volatility models GARCH

- After the success of the ARCH model several extensions have been proposed by researchers.
- Probably the most popular extension is the GARCH( $p,q$ ) introduced by Bollerslev (1986).
- GARCH = generalized autoregressive conditional heteroscedasticity

### GARCH-type volatility models GARCH

- The GARCH(1,1) model is specified as

$$y_t = \omega + \varepsilon_t$$

$$\varepsilon_t = \sigma_t u_t, \quad u_t \sim N(0,1)$$

$$(\sigma_t)^2 = \alpha_0 + \alpha_1 (\varepsilon_{t-1})^2 + \beta_1 (\sigma_{t-1})^2$$

where  $\alpha_i > 0$  for  $i=0,1$  and  $\beta_1 > 0$  to ensure the positive value of volatility and  $\omega$  is a real number.

### The EWMA model

- A special case of the GARCH(1,1) model used in practice is the **exponentially weighted moving average** (EWMA) model.
- The EWMA model has been introduced by J.P. Morgan investment bank in 1996.

### The EWMA model

- The EWMA model is formulated as follows:

$$y_t = \omega + \varepsilon_t$$

$$\varepsilon_t = \sigma_t u_t, \quad u_t \sim N(0,1)$$

$$(\sigma_t)^2 = (1-\lambda) (\varepsilon_{t-1})^2 + \lambda (\sigma_{t-1})^2$$

with  $\lambda=0.94$

- Notice that one does not have to estimate the parameters of the volatility equation. The  $\lambda$  is set to 0.94.

### GARCH-type volatility models

#### GARCH

- The GARCH( $p,q$ ) is formulated as follows:

$$y_t = \omega + \varepsilon_t$$

$$\varepsilon_t = \sigma_t u_t, \quad u_t \sim N(0,1)$$

$$(\sigma_t)^2 = \alpha_0 + \alpha_1 (\varepsilon_{t-1})^2 + \alpha_2 (\varepsilon_{t-2})^2 + \dots + \alpha_q (\varepsilon_{t-q})^2 + \beta_1 (\sigma_{t-1})^2 + \beta_2 (\sigma_{t-2})^2 + \dots + \beta_p (\sigma_{t-p})^2$$

where  $\alpha_i > 0$  for  $i=0, \dots, q$  and  $\beta_j > 0$  for  $j=1, \dots, p$  to ensure the positive value of volatility and  $\omega$  is a real number.

### GARCH-type volatility models

#### GARCH

##### Stationarity

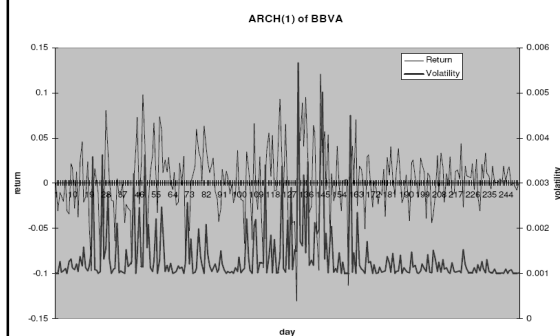
- The GARCH(1,1) model is stationary if  $\alpha_1 + \beta_1 < 1$ .
- The GARCH( $p,q$ ) model is stationary if

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$$

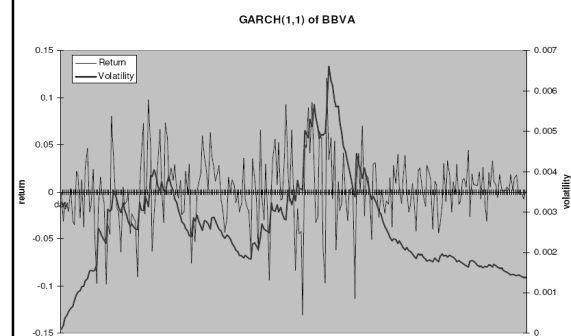
### EXAMPLE: Volatility of BBVA

- We present an example of the volatility estimates for return data of the BBVA stock during one year using the ARCH(1) and GARCH(1,1) models.

### EXAMPLE: ARCH(1) of BBVA



### EXAMPLE: GARCH (1,1) of BBVA





## 1(c) EGARCH MODEL

### GARCH-type volatility models EGARCH

- A further modification of the GARCH model is the exponential-GARCH or EGARCH developed by Nelson (1991).
- The EGARCH model allows for asymmetry in volatility.
- The EGARCH(1,1,1) is formulated as follows:

### GARCH-type volatility models EGARCH(1,1,1)

$$y_t = \omega + \epsilon_t$$

$$\epsilon_t = \sigma_t u_t \quad u_t \sim N(0, 1) \text{ i.i.d}$$

$$\ln \sigma_t^2 = \alpha_0 + \alpha_1 \left| \frac{\epsilon_{t-1}}{\sigma_{t-1}} \right| + \beta_1 \ln \sigma_{t-1}^2 + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}}$$

where the  $\gamma$  parameter controls for asymmetry.

- The EGARCH parameters are not restricted: they are real numbers.

### GARCH-type volatility models EGARCH

#### Stationarity

- The EGARCH(1,1,1) model is stationary when  $|\beta_1| < 1$ .

## 2. STOCHASTIC VOLATILITY MODEL

### Stochastic volatility (SV) models

- In the GARCH-type models changing volatility is driven by past **squared returns** (ARCH, GARCH) or past **absolute returns** (EGARCH).
- Both these are alternative measures of volatility.

### Stochastic volatility (SV) models

- An alternative possibility for dynamic volatility is the so-called stochastic volatility model.
- In this model we introduce an **innovation term** (or 'error term') into the volatility equation.

### Stochastic volatility (SV) model

- The SV(1) model is formulated as:

$$y_t = \omega + \epsilon_t \quad \epsilon_t = \sigma_t u_t$$

$$u_t \sim N(0, 1) \text{ i.i.d}$$

$$\ln \sigma_t^2 = \alpha_0 + \beta \ln \sigma_{t-1}^2 + v_t$$

$$v_t \sim N(0, \sigma_v^2) \text{ i.i.d}$$

- All SV parameters are real numbers.

### Stochastic volatility (SV) model Stationarity

- The SV(1) model is stationary when  $|\beta| < 1$ .
- The model can be extended to include more lags of volatility.

### Comparison of different volatility models

### Comparison of volatility models

- We have seen several volatility models:
  - ARCH( $q$ ) with  $q=1,2,3,\dots$
  - GARCH( $p,q$ ) with  $p$  and  $q = 1,2,3,\dots$
  - EWMA
  - EGARCH(1,1)
  - OSV(1)
- Each of these models give different estimates of volatility.

### Comparison of volatility models

#### How to choose the best model?

- We have to compare the estimated volatility with the 'true' volatility.
- The problem is that the 'true' volatility is unobservable.
- In practice, we choose a proxy of 'true' volatility.
- A popular proxy of the variance of return is  $(\epsilon_t)^2$

### Comparison of volatility models

- We can follow the next process to measure the quality of a volatility model:
  - (1) Estimate the parameters of the volatility model
  - (2) Compute the estimated variance of returns  $(\sigma_t)^2$
  - (3) Compute the proxy  $(\epsilon_t)^2$
  - (4) Compute the mean squared error (MSE) between  $(\sigma_t)^2$  and  $(\epsilon_t)^2$ :

$$\text{MSE} = \frac{1}{T} \sum_{t=1}^T (\sigma_t^2 - \epsilon_t^2)^2$$

### Comparison of volatility models

- Repeat this procedure for several volatility models to compute the MSE for each of them.
- The best volatility model is the one with the **lowest** MSE.

### Explanatory variables in the volatility equation

#### Explanatory variables in the volatility equation

- In the previous models, we did not include explanatory variables into the volatility equation.
- In practice, it may be useful to consider these variables in the volatility formulation because:
  - They may help to solve to problem of **non-stationarity**.
  - They may **improve the quality of volatility estimation**.

#### Explanatory variables in the volatility equation

- In the followings, we specify the three volatility models with exogenous variables.

##### ● ARCH(1)-X model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \delta X_t$$

#### Explanatory variables in the volatility equation

##### ● GARCH(1,1)-X model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \delta X_t$$

##### ● EGARCH(1,1)-X model:

$$\ln \sigma_t^2 = \alpha_0 + \alpha_1 \left| \frac{\epsilon_{t-1}}{\sigma_{t-1}} \right| + \beta_1 \ln \sigma_{t-1}^2 + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \delta X_t$$

### Explanatory variables in the volatility equation

#### Stationarity:

- In these models, we have the same conditions of stationarity as in the models without explanatory variables.
- The  $X_t$  variable may solve the stationarity problem of the volatility model.

### Explanatory variables in the volatility equation

#### Positive parameters in ARCH, GARCH

- In the ARCH and GARCH models, have the positivity restriction of all parameters, including the  $\delta$  parameter of the explanatory variables.
- This may be problematic in some cases when a negative sign is expected for  $\delta$ .
  - Example: If  $X_t$  is the product price of a firm then decreasing  $X_t$  may increase the firm's risk.

### Explanatory variables in the volatility equation

- In the EGARCH model, there is no sign restriction on  $\delta$ .
- Therefore, when one wants to include additional variables into the dynamic volatility model, it is suggested to use the EGARCH specification as there is no sign restriction in that model.

### Volatility targeting

### Volatility targeting - motivation

- Practitioners in many occasions are interested in predicting future volatility, i.e. volatility targeting.
- This may be helpful for derivatives pricing reasons or portfolio construction purposes.
- The dynamic models of volatility give forecasting formulas for future volatility.

### Volatility targeting in GARCH(1,1)

- We are in the beginning of period  $t$ , i.e. we know  $F_{t-1} = (y_1, \dots, y_{t-1})$
- First, we estimate the parameters of the GARCH(1,1) model.
- Then, we are interested in the forecasts of future volatility for the periods  $t+1, t+2, \dots$ .
- These forecasts we can compute substituting the parameters estimates into the next formula:

### Volatility targeting in GARCH(1,1)

- The  $s$ -step-ahead volatility forecast is:

$$E(\sigma_{t+s}^2 | F_{t-1}) = \sigma^2 + (\alpha_1 + \beta_1)^s (\sigma_t^2 - \sigma^2)$$

- for  $s=1,2,\dots$  future time horizon where

$$F_{t-1} = (y_1, \dots, y_{t-1})$$

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}$$

### Volatility targeting in GARCH(1,1)

- We consider the GARCH(1,1) volatility model because of the simplicity of the forecast formula.
- Nevertheless, before using the next formula one needs to check if the parameter  $\alpha_1$  is significant and the model is stationary, i.e.  $\alpha_1 + \beta_1 < 1$ .

### Estimating volatility for the Black-Scholes model

### Estimating volatility for the Black-Scholes model

- When we price European options using the Black-Scholes (BS) formula, we need an estimate of the volatility of the underlying product,  $\sigma$ .
- The volatility in the BS formula refers to the standard deviation of log returns at the **annual** scale.
- In practice, however, we have price data on the daily frequencies.

### Estimating volatility for the Black-Scholes model

- **How to get the estimate of this annual volatility?**

- There are several ways to proceed:

### Constant volatility

- (1) Assuming that daily volatility is **constant**, we can use the next equation:

$$\sigma_{\text{annual}} = \sigma_{\text{daily}} \sqrt{250}$$

- where we simply rescale the constant daily volatility to the annual, 250-trading day, horizon.
- The daily volatility can be estimated as before in Excel computing the standard deviation of returns.

### Changing volatility

- However, as we have seen before, volatility is not constant and the dynamic volatility models explain its changing behavior.
- We use the GARCH(1,1) model to compute changing annual volatility.
- We proceed as follows:

### Changing volatility

- (2a) Estimate a GARCH(1,1) model to past daily log returns.
- (2b) Make sure that the parameters are significant and that the GARCH(1,1) is stationary.
- (2c) Forecast volatility for the next annual period (for  $s=1, \dots, T$ ) by the next formula:

$$E(\sigma_{t+s}^2 | F_{t-1}) = \sigma^2 + (\alpha_1 + \beta_1)^s (\sigma_t^2 - \sigma^2)$$

### Changing volatility

- (2d) Compute the sum of forecasted variances over the next 1-year period:

$$\sigma_{\text{annual}}^2 = \sum_{s=1}^T E(\sigma_{t+s}^2 | F_{t-1})$$

- (2e) Take the square-root of the previous sum:

$$\sigma_{\text{annual}} = \sqrt{\sigma_{\text{annual}}^2}$$

### Black-Scholes formula

- After estimating the annual volatility of log return, we can substitute it into the BS formula to compute the option price.
- In the next slides, we show the BS formulas for call and put options:

### BLACK-SCHOLES FORMULA

- The BS price of a European call option:

$$c = S_0 N(d_1) - X \exp(-rT) N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S_0}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

### BLACK-SCHOLES FORMULA

#### Notation:

- $S_0$ : price of the underlying asset at  $t=0$ .
- $X$ : strike price of the option.
- $r$ : risk-free rate.
- $T$ : time-to-expiration of the option.
- $\sigma$ : standard deviation of the return of the underlying asset ("volatility") at annual scale.
- $N(\cdot)$ : distribution function of the standard normal distribution.

### BLACK-SCHOLES FORMULA

- The BS price of a European put option:

$$p = X \exp(-rT)N(-d_2) - S_0N(-d_1)$$

$$d_1 = \frac{\ln\left(\frac{S_0}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

### BLACK-SCHOLES FORMULA

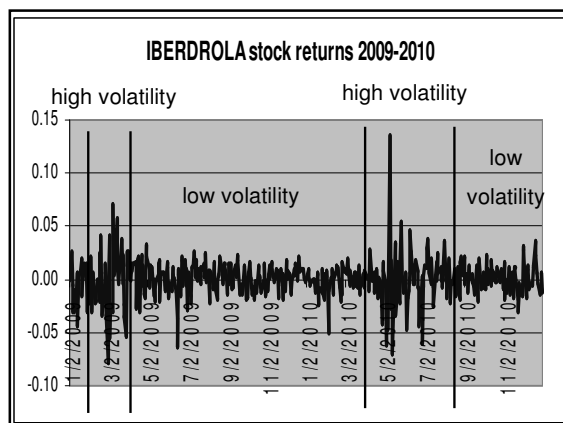
#### Notation:

- $S_0$ : price of the underlying asset at  $t=0$ .
- $X$ : strike price of the option.
- $r$ : risk-free rate.
- $T$ : time-to-expiration of the option.
- $\sigma$ : standard deviation of the return of the underlying asset ("volatility") at annual scale.
- $N(\cdot)$ : distribution function of the standard normal distribution.

## Jumps in security prices: Markov switching models

### Markov switching models - Motivation

- In the financial markets it is empirically observed that periods of higher volatility (for example around financial crisis periods) and periods of lower volatility (during 'normal' economic periods) change one another over time.
- For example, the next figure shows the stock returns of Iberdrola company during 2009 and 2010.



### Markov switching models

- The previous figure motivates econometric models, which assign high volatility during certain periods and low volatility during another periods.
- The models are called the **Markov switching** (MS) or **regime switching** models.
- In the MS models, the parameters of the models switch their values over time.

### Markov switching models

- In this topic, we see a simple formulation for the MS model, which capture switches in the mean as well as in the volatility.
- **The model to be considered involves switches between two alternative regimes denoted by  $s_t$ .**
- The variable of the regimes,  $s_t$  may take two alternative values:  

$$s_t = 0 \text{ or } s_t = 1$$
- $s_t$  is **unobservable** for the economist

### Markov switching models

- We consider a linear regression model with switching parameters, i.e. **the parameters of the model are functions of  $s_t$**
- In other words, the parameters of the model are **time dependent**.



### Markov switching models

- The MS model is formulated as follows:

$$y_t = c(s_t) + \psi(s_t)u_t$$

- where the error term has normal distribution:

$$u_t \sim N[0, 1]$$

### Markov switching models

$$s_t \in \{0, 1\}$$

- indicates the regime at time  $t$ , which forms a Markov process with a 2x2 transition probability matrix

$$P = \{\eta_{ij}\}$$

- for  $i, j=0, 1$  and is defined as follows:

### Markov switching models

$$\Pr[s_t = 0 | s_{t-1} = 0] = \eta_{00}$$

$$\Pr[s_t = 1 | s_{t-1} = 0] = \eta_{10}$$

$$\Pr[s_t = 0 | s_{t-1} = 1] = \eta_{01}$$

$$\Pr[s_t = 1 | s_{t-1} = 1] = \eta_{11}$$

### Markov switching models

- Notice that

$$\eta_{00} + \eta_{10} = 1$$

$$\eta_{11} + \eta_{01} = 1$$

- Therefore, it is enough to estimate  $\eta_{00}$  and  $\eta_{11}$  because the other two parameters of  $P$  can be expressed from these two equations.

### Markov switching models

- In the MS model,

$$y_t = c(s_t) + \psi(s_t)u_t$$

- the following parameters are included:
- $c(0)$ ,  $c(1)$ ,  $\psi(0)$ ,  $\psi(1)$ ,  $\eta_{00}$  and  $\eta_{11}$
- $c(0)$  and  $c(1)$  capture the **mean** of the return
- $\psi(0)$  and  $\psi(1)$  measure the **volatility** of the return

### Estimation of the Markov switching model

### Estimation

- The MS model is estimated by the maximum likelihood technique.
- The likelihood function is given by:

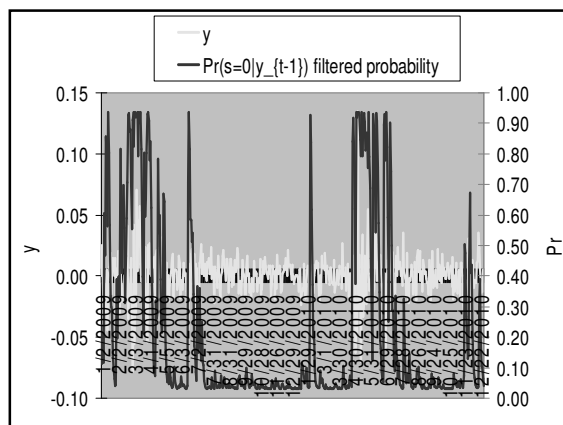
$$L = \prod_{t=1}^T \pi_{0t} f(y_t | Y_{t-1}, s_t = 0) + \pi_{1t} f(y_t | Y_{t-1}, s_t = 1)$$

- where  $\pi_{jt} = \Pr[s_t = j | Y_{t-1}]$
- is called the 'filtered probability' of regime  $j$ .

### Estimation

- The parameter estimates are obtained by maximizing  $\ln L$  in the parameters.
- See the Markov - notes.pdf file for the description of the computation of the likelihood function.

- See the estimated probability of regime 0 for Iberdrola on the following figure:



### Estimation of the mean

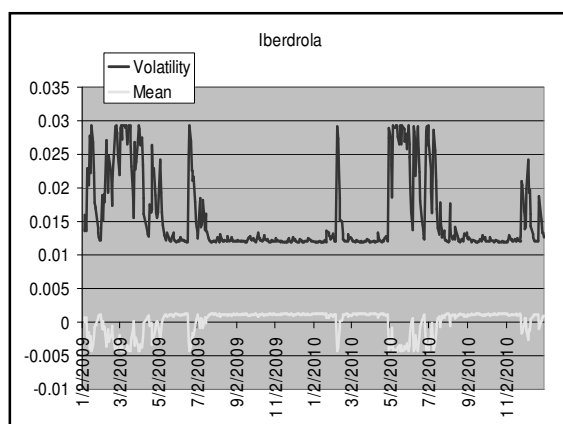
- The time dependent **mean** is estimated as

$$\hat{c}_t = \pi_{0t} \hat{c}(0) + \pi_{1t} \hat{c}(1)$$

- The time dependent **volatility** is estimated as

$$\hat{\psi}_t = \pi_{0t} \hat{\psi}(0) + \pi_{1t} \hat{\psi}(1)$$

- See the estimated mean and volatility for Iberdrola on the next figure:



### Practice

- Compute the log likelihood of the data.
- Then, maximize  $\ln L$  by Solver to find the parameter estimates.
- Graph the probability of the first regime  $j=0$ , i.e.

$$\pi_{jt} = \Pr[s_t = j | Y_{t-1}]$$

## Extreme values in security returns: value at risk

### Value at risk - Motivation

- **Extreme price movements in the financial markets are rare, but important.**
- The stock market crash on Wall Street in October 1987 and other big financial crises such as the Long Term Capital Management (LTCM) have attracted a great deal of attention among practitioners and researchers.
- Some people even called for government regulations on financial risk management.

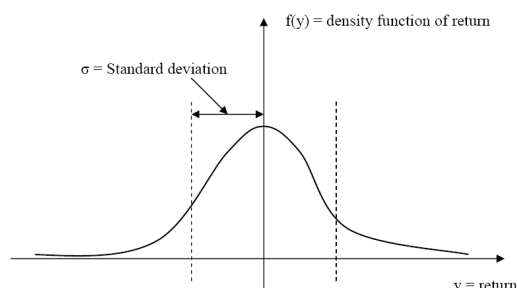
### Value at risk - Motivation

- The recent 2008/09 financial and economic crisis have generated further discussions on market risk.
- In this section, we study the impact of unlikely **extreme events** on portfolio return.
- During the past two decades, **value at risk (VAR)** has become a widely used measure of market risk in risk management.

### Value at risk - Motivation

- In the previous sections, we defined financial risk as the standard deviation of the rate of return on financial assets.
- In this definition, we use a particular characteristic of uncertain returns: the **standard deviation** that represents the overall variability of financial asset prices.
- In the following figure, we demonstrate the standard deviation risk measure for the example of normal distribution of returns:

### Value at risk - Motivation



### Value at risk - Motivation

An alternative measure of financial risk is the so-called **value at risk (VAR)**.

- The value at risk is used by financial institutions in order to measure the risk of their portfolios.
- Banks in several countries are **legally obliged** to compute and report the VAR of their portfolios.
- What is value at risk?

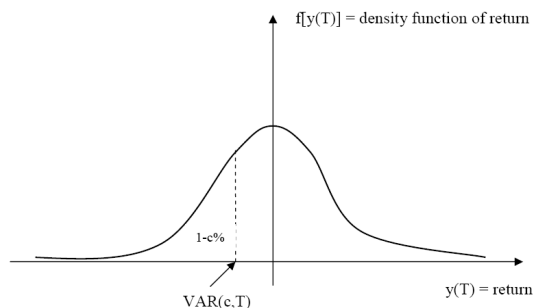
### Value at risk - definition

- **VAR answers the following question:**
- With probability  $p$ , how much the investor is going to loose on his portfolio during the next  $T$  days?
- **Answer:**
- The investor will loose at least  $\text{VAR}(1-p, T) = \text{VAR}(c, T)$  where  $p$  denotes the probability,  $c$  the *confidence level* and  $T$  the time duration of the investment.

### Value at risk - definition

- Graphically, we can define the  $\text{VAR}(c, T)$  corresponding to confidence level  $c$  and time horizon  $T$  as follows:

### Value at risk - definition



### Value at risk - definition

- There are two elements in the definition of VAR:
- (1) **Confidence level,  $c$ :** This determines the  $p=1-c$  probability of the extreme loss.  
Ocommonly used values for  $c$  are 90%, 95%, 97.5%, 99%
- (2) **Target horizon,  $T$ :** This determines the time horizon of the return of interest  
Ofor example,  $T = 1$  day, 1 week, 1 month, etc.

### Value at risk - definition

- For the different confidence levels,  $c$  (i.e. probabilities,  $p=1-c$ ) and time horizons,  $T$  the VAR estimate will be different.

## Computing VAR

## Computing the value at risk

- We provide three alternative ways to compute the VAR of the portfolio return:

1. **Historical VAR**
2. **Delta-normal VAR**
3. **Monte Carlo VAR**

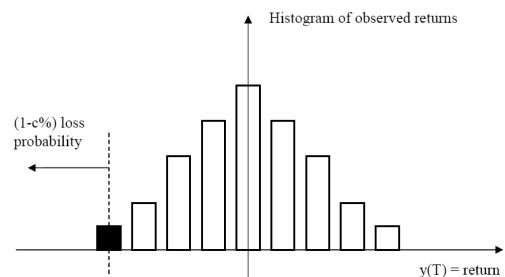
## 1. HISTORICAL VAR

### HISTORICAL VAR

- In the historical VAR, first, collect data on **historical returns**.
- Second, **order** returns from the lowest to the highest in an increasing order.
- Third, starting with the lowest return, **count** the number of returns until we get to the  $p=1-c$  portion of total number of observations.
- In Excel, the counting done using the PERCENTIL(data set,  $1-c$ ) function:
- $\text{VAR}(c, T) = \text{PERCENTIL}(\text{data set}, 1-c)$

### HISTORICAL VAR

- Graphically, we can illustrate this on the next figure:



### HISTORICAL VAR

- The **advantage** of the historical VAR is that we do not use any assumption regarding the distribution of returns.
- We use a historical data set to determine VAR without any additional assumptions on the return distribution, i.e. we do not need to estimate the parameters of the return distribution.

### HISTORICAL VAR

- The **disadvantage** of the historical VAR is that we only use past observations to infer the future return distribution.
- The data set of past returns may be small and/or non-representative for the inference of the  $p=1-c$  quantile of future returns.

## 2. DELTA-NORMAL VAR

### DELTA-NORMAL VAR

- We are in day  $t=0$  (present) and we are interested in computing VAR of portfolio invested over a time horizon of  $T$  days.
- Denote the log return of the investment by:

$$y(T, 0) = \ln \left( \frac{p_T}{p_0} \right)$$

- $y(T, 0)$  is the log return obtained between time  $t=0$  and time  $t=T$ .

### DELTA-NORMAL VAR

- In the delta-normal approach, we assume that log return is normally distributed:
- $y(T, 0) \sim N[\mu(T), \sigma^2(T)]$

### DELTA-NORMAL VAR

- This assumption allows us to obtain the following formula for  $\text{VAR}(c, T)$ :

$$\text{VAR}(c, T) = \mu(T) + \sigma(T) Z_{1-c}$$

- where  $Z_{1-c}$  is the "quantile  $1-c$ " of the standard normal distribution  $N(0, 1)$ .

### DELTA-NORMAL VAR

- The following table gives you some important values of  $Z_{1-c}$  of the standard normal distribution, which are used in practice:

Level of confidence = c	1-c	Quantile of $N(0,1) = Z_{1-c}$
99%	1%	-2.32635
97.5%	2.5%	-1.95996
95%	5%	-1.64485
90%	10%	-1.28155

### DELTA-NORMAL VAR

- The main **advantage** of the delta-normal VAR is that it is computationally easy to get the VAR estimate:
- We only have to substitute the estimated parameters  $\mu(T)$ ,  $\sigma(T)$  and  $Z_{1-c}$  into the previous VAR formula.

### DELTA-NORMAL VAR

- The **disadvantage** of the delta-normal VAR is that we need to estimate  $\mu(T)$  and  $\sigma(T)$  using historical returns with time horizon  $T$ .
- (1) Thus, the delta-normal VAR also uses historical data to estimate the parameters.
- (2) The estimation of  $\mu(T)$  and  $\sigma(T)$  requires a large data set of  $y(T,0)$  that may not be available.

### DELTA-NORMAL VAR

#### Example

- Imagine that we compute 30-days delta-normal VAR.
- For this, we estimate  $\mu(T)$  and  $\sigma(T)$  using a sample of 120 observations of the past.
- This means a 120 x 30-days time span, which is approximately 10 years of past data.
- This may not be available for us.

### DELTA-NORMAL VAR

- In practice, it may happen that the time horizon of the returns for which one has sufficient sample size is different from the time horizon of the  $\text{VAR}(c, T)$ .
- In particular, it is possible that daily log returns are observed  
 $y(t, t-1)$  for  $t=1, 2, 3, \dots, T$
- How to compute  $\text{VAR}(c, T)$  in this case?

### DELTA-NORMAL VAR

- The formula of  $\text{VAR}(c, T)$  for the case when  $y(t, t-1)$  daily log returns are observed is the following:

$$\text{VAR}(c, T) = \mu(T) + \sigma(T)Z_{1-c} = T\mu(1) + \sqrt{T}\sigma(1)Z_{1-c}$$

- where  $\mu(1)$  and  $\sigma(1)$  denote the expectation and standard deviation of daily log return, respectively.

## 3. MONTE-CARLO VAR

### MONTE CARLO (MC) VAR

- The third method of VAR computation is based on Monte Carlo **simulation** of future returns.
- First, similarly to the delta-normal VAR, we assume a specific distribution of returns.
- Then, we simulate a large number of returns from this distribution.
- Each simulation is interpreted as a possible realization of future return.

### MONTE CARLO (MC) VAR

- For example, suppose that we compute the 1-day VAR( $c,1$ ) using MC method to estimate the largest possible loss during tomorrow.
- To do this, we simulate thousands of realizations from  $y(1,0)$ .
- Each of these simulations is interpreted as a possible return of tomorrow.

### MONTE CARLO (MC) VAR

- Suppose the risk manager estimates the mean and volatility of returns of the portfolio and wants to simulate future returns from
- $y \sim N(\mu, \sigma^2)$ .
- **How to simulate  $y \sim N(\mu, \sigma^2)$  returns in Excel?**

### MONTE CARLO (MC) VAR

- **Simulate in three steps:**
  - (1) Simulate random numbers between 0 and 1 denoted by  $U$  by the RAND() function.
  - (2) Transform all  $U$  to  $N(0,1)$  random numbers by  $Z = \text{NORMSINV}(U)$ .
  - (3) Compute  $y = \mu + \sigma Z$  for all simulations.
- Then,  $y \sim N(\mu, \sigma^2)$ .

### MONTE CARLO (MC) VAR

#### **Simulation and VAR stability:**

- One can re-simulate random return data in Excel by the F9 button.
- The VAR estimate changes in every simulation because of the new random returns.
- One can make the VAR estimate more stable by increasing the number of simulations.

### MONTE CARLO (MC) VAR

- From the returns simulated, the VAR is determined numerically by the procedure used in the historical VAR computation:
- VAR( $c, T$ ) = PERCENTIL(data set,  $1-c$ ).

### MONTE CARLO (MC) VAR

- The **advantage** of the MC method is that we can simulate a large number of returns for the same time horizon when we compute VAR.
- Thus, we are not limited by a small sample size.
- Recall that for the historical VAR, we observe returns only for a limited historical time period – so we are limited by the sample size.



### MONTE CARLO (MC) VAR

- The **disadvantage** of the MC method is that we need to assume a distribution of the returns and we need to estimate properly the parameters of that distribution using historical data on returns.
- Therefore, we can make error by choosing a wrong distribution and/or estimating incorrectly the parameters of the distribution.

### A NOTE ON THE DISTRIBUTION OF ASSET RETURNS

– ‘fat tails’ or excess kurtosis of returns

#### A note on the distribution of asset returns

- When computing VAR it is frequently assumed that the distribution of the rate of return is a normal distribution.
- The delta-normal VAR makes this assumption and we frequently simulate from an assumed normal distribution in the MC VAR as well.

#### A note on the distribution of asset returns

- The main reasons of the assumption of normality is computational convenience.

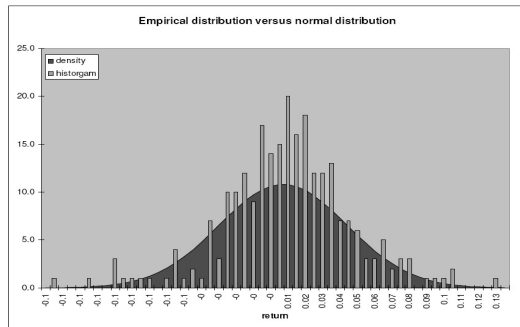
#### A note on the distribution of asset returns

- However, there is empirical evidence that the distribution of returns is not normal.
- The true distribution of returns has so-called “**fat tails**” or **excess kurtosis**.
- This means that extreme observations occur more probably than it is explained by the normal distribution.

#### A note on the distribution of asset returns

- This issue is important from the risk management point of view, where we are especially interested in the proper modelling of the extreme negative observations.
- The following figure shows this phenomenon for the BBVA stock daily returns for data collected over a one-year period:

### A note on the distribution of asset returns



### A note on the distribution of asset returns

- In practice, there are alternative solutions for the 'fat tails' problem:
- (1) Use distributions, which fit well to real data because they exhibit fat tails – **'fat tailed' distributions**.
- (2) Use the standard non-fat tailed distributions like the normal distribution but estimate every day a new value for its volatility – **use dynamic models of volatility**.

#### (1) Some 'fat-tailed' distributions

##### (1) Student-t distribution

○ A symmetric distribution frequently used in statistics, one of its properties is that it has fat-tails.

##### (2) Lévy distribution

○ Lévy distribution is a generalization of the normal distribution where an additional parameter controls for fat-tails.

##### (3) Generalized error distribution

○ This is a distribution with zero mean and variance with an additional parameter controlling for fat-tails.

#### (2) Update return volatility

- For example, we may assume that log returns are normally distributed with time dependent parameters  $\mu_t$  and  $\sigma_t$ .
- Then, we can use a dynamic volatility model – for example GARCH – the update  $\sigma_t$ .
- This way when the value of  $\sigma_t$  is high, for example, we can capture extreme price deviations of the portfolio.

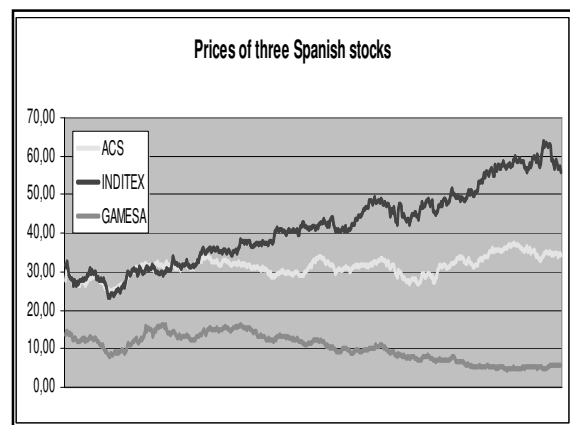
## Portfolio selection based on cointegration

### Motivation – financial data used

- In the previous topics, the financial data applied in the models have been the **return** obtained on financial assets.
- In this section, we use **price levels** in the econometric models.

### Motivation – random walk of financial asset prices

- Collecting price data on most financial assets traded on exchanges and plotting the evolution of these data we have figures like this:



### Motivation – random walk of financial asset prices

- A typical characteristic of the financial price time series is that they follow a so-called random walk.
- Generally speaking, this means that these prices are not “forecastable” based on past information and they do not follow time trends or move around a constant value that could determine the approximate location of future prices.

### Motivation – random walk of financial asset prices

- The random walk process of price  $s_t$  can be characterized by the following equation:

$$s_t = c + s_{t-1} + u_t$$

with  $E[u_t] = 0$  and  $\text{Var}[u_t] = \sigma^2$

- The equation includes two parameters:  $c$  is the constant term and  $\sigma^2$  is the variance of the error term,  $u_t$

Motivation – random walk of financial asset prices

- In statistics, the random walk process is often called **unit root process**.
- These processes have very uncertain behaviour, they are **non-stationary**, and are **non-mean-reverting processes**.
- Therefore, it is not useful for practitioners to work with them directly.

Motivation – cointegration

- However, there is a field of econometrics, called cointegration, where these processes are transformed in a way that may be useful for practitioners.
- The main idea is as follows:

Motivation – cointegration

- Suppose that we have  $N$  unit root processes, say asset prices of a portfolio with  $N$  assets, where the prices of all assets in the portfolio move according to a random walk.
- **These unit root processes are said to be cointegrated if a linear combination of their processes forms a covariance stationary process.**

Motivation – cointegration

- Two comments on this definition related to
  - (1) covariance stationarity and
  - (2) linear combination:

Motivation – (1) stationarity

When a process is “covariance stationary”?

- It is covariance stationary if it is stable in the following sense:
- (1) the expected value of the price is constant, i.e.  $E[s_t] = \alpha$  for all  $t$ .
- (2) the variance of the price is constant, i.e.  $\text{Var}[s_t] = \sigma^2$  for all  $t$ .
- (3) the covariance between prices of different time periods does not depend on  $t$  in the following sense:  $\text{cov}[s_t, s_{t-j}] = \gamma_j$

Motivation – (1) stationarity

- An important and useful implication of covariance stationarity is the mean reverting property of the process.
- This means that the process is moving around the constant mean,  $\alpha$  and it is always reverting to it, i.e. it cannot go very far away from its mean.
- See the following example.

### Motivation – (1) stationarity

- Example: Let us consider a particular and simple stationary process, the first-order autoregressive, AR(1), process.
- The AR(1) process is formulated as follows:

$$s_t = c + \phi s_{t-1} + u_t$$

with  $E[u_t] = 0$ ,  $\text{Var}[u_t] = \sigma^2$  and  $|\phi| < 1$

### Motivation – (1) stationarity

- When  $|\phi| < 1$ , then this process is stationary and therefore, it is mean reverting.
- The mean of the AR(1) process is computed as:

$$E[s_t] = \frac{c}{1 - \phi}$$

### Motivation – (2) linear combination

- The linear combination of  $N$  variables,  $s_{it}$  with  $i=1, \dots, N$  with respect to the coefficients  $\beta_i$  with  $i=1, \dots, N$  is the following:

$$p_t = \sum_{i=1}^N \beta_i s_{it}$$

- The coefficients  $\beta_i$  are constant parameters.
- In the framework of cointegration, the vector  $(\beta_1, \dots, \beta_N)$  is called “cointegrating vector”.

### Motivation – (2) linear combination

- Notice that the previous formula is the way we compute the price of a portfolio of the  $N$  assets when quantities  $\beta_i$  are purchased ( $\beta_i > 0$ ) or short sold ( $\beta_i < 0$ ) from each asset.
- This idea gives us the practical motivation for considering the idea of cointegration for portfolio construction reasons.

### Motivation – summary

- The previous slides can be summarized as follows.
- Given that we have a set of assets whose prices move according to a random walk process, if they are cointegrated then it is possible to create a portfolio with a covariance stationary, and therefore, a mean reverting value.

### Motivation – summary

- As the value of this portfolio is reverting to its mean always, the investor can forecast the future tendency of its value.
- This means that if the current value of the portfolio is lower than its mean, then it makes sense to create a long position in the portfolio.
- On the other hand, if the current value of the portfolio is higher than its mean, then it makes sense to create a short position in the portfolio.

### Motivation – summary

- A “small” problem with these arguments:
- Maybe the value of the cointegrating vector are not valid in the future, i.e. maybe the value of the portfolio created by the cointegrating vector estimated by using past price data does not yield a stationary process for the future.

### The steps of the portfolio construction

- (1) Test the unit root null hypothesis for all assets in the portfolio.
- (2) Considering only the assets whose prices form unit root price processes,
  - (2a) test for the existence of cointegration among the price processes
  - (2b) determine the number of cointegrating vectors (there may be more than one in general)
  - (2c) estimate the components of the cointegrating vector

### The steps of the portfolio construction

- (3) Knowing the initial value of your capital to be invested, rescale the value of each cointegrating vector in such a way that the total value of your portfolio equals to the cash you have.
- (4) Choose one of the cointegrating vectors that you will use to construct the portfolio.
- (5) Take your position in the selected portfolio.
- (6) When the portfolio price process is near the mean, realize the profit by liquidating the portfolio.

### (1) Unit root test of the price processes

- There are many different unit root tests available.
- In this topic, we use the most common one, the Augmented Dickey-Fuller (ADF, Dickey and Fuller, 1979) test.

### (1) Unit root test of the price processes

- The test we use has the following null and alternative hypotheses:
  - $H_0$ :  $s_t$  is a unit root process
  - $H_1$ :  $s_t$  is not a unit root process

### (2) Cointegration test

- (2a)-(2c) we can obtain as follows:
  - Quick / Group Statistics / Johansen Cointegration test...
  - Write in a list all assets whose process is unit root according to the ADF test
  - In the appearing window, do not change anything, just push OK.
  - The window will show us all results of (2).

## (2) Cointegration test

- Interpretation of the results window of the Johansen cointegration test:
- The first two windows on "Unrestricted cointegration rank test" give information on the existence of cointegrating relationship and the number of cointegrating vectors.
- The two tables are two alternative test statistics for the same test.
  - (1) Trace statistic
  - (2) Maximum eigenvalue statistic

## (2) Cointegration test

- The conclusions derived from these two tables may not be the same.
- When this happens, it is recommended to use the conclusions derived from the maximum eigenvalue test statistic (i.e. the second table).

## (2) Cointegration test

- How to interpret these tables?
- First, we have to look at the last column where the p-values are reported.
- (2a) Existence of cointegration:
  - If none of the p-values is less than 0.05 then there is no cointegrating relationship among the price processes;
  - If there is/are p-value(s) less than 0.05 then there is cointegrating relationship among the variables.

## (2) Cointegration test

- (2b) Number of cointegrating vectors:
  - If there is one p-value (next to "None") which is less than 0.05 then there are 2 cointegrating vectors;
  - If there are two p-values (next to "None" and "At most 1") which are less than 0.05 then there are 3 cointegrating vectors; etc.

## (2) Cointegration test

- (2c) Components of the cointegrating vector:
  - After determining the number of cointegrating vectors, go down to the table next to "x Cointegrating Equations", where x is the number of cointegrating vectors.
  - The cointegrating vector(s) can be seen in the table "Normalized cointegrating coefficients".

## (3)-(5) Normalization of the cointegrating vector & portfolio selection

- (3) Knowing the initial value of your capital to be invested, rescale the value of each cointegrating vector in such a way that the total value of your portfolio equals to the cash you have.
- (4) Choose one of the cointegrating vectors that you will use to construct the portfolio:
  - Choose the one with the largest absolute difference from the mean of the portfolio
- (5) Take your position in the selected portfolio.